

Can Europe's dream now come true?

If the disappointments of the past few weeks have tarnished the hope of winning ratification of the Maastricht Treaty, why not confess that the treaty is far from perfect and renegotiate it?

THE first of January next year is meant to be a red-letter day in the European calendar. To mark the occasion, the European Commission is trying to organize a network of mutually visible bonfires spanning its territory, all of them to be lit at 2400 hours on 31 December. Why all the fuss? Because that is the instant chosen in 1986 when the 12 states of the European Communities (EC) would become the authentic 'single market' promised by the Treaty of Rome a quarter of a century earlier, but not fully delivered even now. But the rejoicing in January will be neither as general nor as full-throated as the commission hopes. On the contrary, there is now a serious danger that the dream of a Europe united around the nucleus of the EC will have gone sour before the time comes to set the bonfires alight.

That is to say the unspeakable, but the events of the past few weeks allow no choice. Europe is no longer a place on the threshold of a bright new future, but instead seems headed towards the patchwork of competitive chauvinism that has dominated European history since the beginning of the Holy Roman Empire. Last weekend's rally in Berlin, organized by the German president as a demonstration of German willingness to offer asylum to people persecuted elsewhere, has served chiefly to demonstrate the anger of those who believe differently. That coincided with the French government's obduracy over the GATT negotiations and followed by only a few days the British government's shabby manoeuvring to persuade the House of Commons that the Maastricht Treaty should at least be discussed.

Recession

What has gone wrong? First, the boomtime of the 1980s has given way to recession. People generally are less comfortable with the prospect of further change. Britain and Italy have special economic difficulties of their own making, while the high cost of reunification threatens recession in Germany and casts an economic pall over the rest of Europe.

Second, Europe's traditional way of making progress has backfired with the treaty signed last December at Maastricht. For 30 years, the EC has made decisions by agreeing on some recipe for change and on a seemingly impossibly tight deadline to put change into effect. But Denmark voted down the treaty in a referendum at the end of June, France won the narrowest majority for the treaty in September, Britain will take the best part of a year from now to ratify it (if at all) and even Germany has said it will not abandon the Deutschmark for a common currency without

a positive decision by the Bundestag later in the decade.

Embarrassment

The Maastricht Treaty has become both a talisman of European unity and a general embarrassment. Its chief provision, the creation of a common European currency, is a logical necessity for a unified trading block, but the supranational monetary institutions necessary for the management of a common currency have offended many people's sense of their countries' sovereignty, even national identity. Because a unified trading block could not work efficiently without a common currency, the case for Maastricht in some shape or form is overwhelming.

The embarrassment has several faces. Even the central provisions of the treaty on the common currency are poorly thought out. They require, for example, a 'convergence' of governments' economic performance in matters such as the budget deficit and the inflation rate, but fail to specify what happens if economic convergence does not come about. And while the treaty provides for the creation of a fund to smooth out economic disparities within what it calls the "European Union", the principles on which this substantial fund would be distributed are only sketchily referred to.

The chapter on research is a good illustration of the defects of the treaty. Its purpose is given simply as the strengthening of the "scientific and technological bases of Community industry". Collaboration between research organizations is especially virtuous. The treaty would also institutionalize the present system of 'framework' programmes by which the commission's research budget, fixed for three to five years in advance, is spent on hastily recruited research applications without much benefit from what might be called peer review. Neither those who drafted the treaty nor those who signed it appear to have given thought to the questions whether Europe's research should be as tightly hitched to European industry's wagon, whether collaboration remains intrinsically virtuous in a truly single market and whether the thrust of successive framework programmes should be as much influenced as hitherto by the whims of successive research commissioners.

The underlying weakness of the Maastricht document is that it is neither a ringing declaration of principles along the lines of the US Constitution nor a legal contract among the signatories, but a pragmatic compromise between the two kinds of documents. Instead of clarity, its drafters have aimed at ambiguity. Does "ever closer union among the



peoples of Europe" imply a federal state? Federalists say that it does, their opponents otherwise.

And what is to be made of the disjunction between the treaty's declaration that "a common foreign and security policy is hereby established" and the exceedingly tentative arrangements for reaching such a state of grace spelled out in later paragraphs? The declaration is meant to please the enthusiasts, of course, while the recitation of what is not yet possible should comfort the sceptics. Europe's failure even to form a common view of the tragedy of Yugoslavia shows that the sceptics, for the time being, are on safe ground.

In an ideal world, the future of European collaboration would not have been allowed to hang in this way on the future of such a tawdry treaty. It would have helped if the drafters had given themselves more time, or if their political masters had been able to spend more than four days at Maastricht last year. It would have been even better if all concerned had given serious thought to questions entirely overlooked in the treaty — how to avoid misunderstandings between Eurocrats and their electorates (typified by the row about the treaty itself), the balance to be struck between the rush (now postponed) for economic growth and geographical enlargement and economic relations with the larger outside world, both prosperous and poor. But the world is not ideal. Maastricht may be a poor treaty, but it is the only one on offer. Its supporters are right to say that, if it fails, it will not be easy to prevent the dissipation of the momentum that has so far successfully sustained the closer collaboration of a dozen European states.

The British government, by rotation now the holder of the presidency of Europe, is plainly hoping that the summit meeting at Edinburgh next month will help a little to that end. If it proves possible to hammer out an understanding with Denmark that might survive a second referendum, might not the same compromise also placate its own sceptics? But that is a poor way of doing serious international business. It would be better simply to renegotiate Maastricht. People in other walks of life often make mistakes, acknowledge them as such and then make them good. Are not politicians at least as vulnerable? And it would be a great misfortune to let Europe fall apart simply because statesmen are unwilling to confess that they signed a piece of paper that does not represent their intentions. □

Clinton on science

The election of Bill Clinton as US president is good news for fetal tissue research and for AIDS.

ALTHOUGH it is too early to say with any certainty what US President-elect Bill Clinton will do about science policy in general (see page 95), his election clearly augurs a change in policy on the use of fetal tissue in human research. With a stroke of the pen, Clinton can rescind the Bush administration's ban on fetal tissue studies and he is expected to do just that within days of taking office. It will be a good decision.

The Bush administration's objections are grounded in the

belief (for which there is no evidence) that clinical research with fetal tissue will encourage women to have abortions. A presidentially appointed panel of scientists and philosophers that included Bernadine Healy, the Republican director of the National Institutes of Health (NIH), concluded two years ago that, with the proper safeguards, fetal tissue research should receive federal support. Ironically, it is the Democrat Clinton who is poised to accept the advice of the Bush panel.

Clinton is also expected to recognize AIDS research in a way the Bush White House never has, with the appointment of a so-called "AIDS tsar" to coordinate research throughout the government—including the NIH, the Centers for Disease Control and the Department of Defense which has concentrated most of its AIDS funds in the US Army. Although the US government is spending \$2 billion on AIDS research, continuing complaints that the effort lacks unified direction have merit. If Clinton names an AIDS tsar (and he should), he should locate the office somewhere in the White House rather than in one of the departments of government. Further, he should select someone with the appropriate scientific knowledge but who is not running an active research laboratory. Credibility and independence require that a White House AIDS coordinator should not compete for funds or credit with those labouring in the laboratory. □

Telephone roundabouts

Cellular telephony seems likely to turn the US telecommunications business upside down.

TECHNOLOGY has a nasty habit of making a monkey of the best laid plans. That is what may be happening to the careful arrangements made in the United States to break up the telephone monopoly known as the Bell System. For five years, federal Judge Harold Green laboured to create some semblance of competition from pieces of the holding company, known as AT&T. His solution, in 1984, seemed sensible enough. The Bell System would be broken up while a company called AT&T would continue to exist to provide long distance and international communications services in competition with then newly licensed companies with names such as MCI and Sprint. Judge Green, believing that the physical ownership of telephone networks gave the baby Bells a *de facto* monopoly, clipped their wings by denying them long-distance business.

That was eight years ago. Now, guess what? AT&T seems to be well placed to recover its old position as the major telephone provider. Last week, Wall Street was in a whirl when told that AT&T plans to buy more than a third of a company called McCaw, which provides cellular telephone services in the United States. McCaw seems to have acquired its 2 million subscribers in the interval since Judge Green's consent decree, and with the help of a technology he had not foreseen. With AT&T's help, it will probably now grow even faster, skimming the cream from the baby Bells to share it with AT&T. The moral? There is none, unless it is that people who expect technology to stand still expect too much. □

Britain prepares to leave European fast-breeder reactor programme

London & Munich. Almost exactly ten years after the United States decided to abandon its efforts to develop a commercial fast breeder reactor, Britain is poised to withdraw from European attempts to achieve the same goal. Its departure could doom the entire project.

Officials at the Department of Trade and Industry were reported last week to be recommending that the British government withdraw its support from the joint European Fast Breeder (EFR) programme, begun in 1984 by Britain, France and Germany. The programme is intended to design and eventually build a reactor demonstrating the commercial feasibility of producing electricity through the action of fast neutrons. The government argues that any responsibility for proceeding with the construction of the reactor should now rest with the nuclear industry.

Official confirmation of Britain's withdrawal was expected to be announced in London this week as part of a package of public spending cuts. Given doubts about proceeding in both France and Germany, such a decision could precipitate the indefinite postponement of any attempt to construct a new fast breeder in Europe and leave Japan as the only country still committed to such a step.

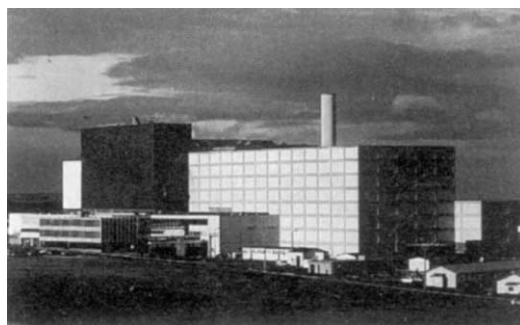
The decision also casts doubt on Britain's reprocessing of spent fuel from its thermal reactors. One of the main justifications for building the Thermal Oxide Reprocessing Plant (THORP) at Sellafield has been the need to obtain plutonium for fast breeders. This justification will disappear if no fast breeders are to be built in Britain in the near future.

The final agreement set out three stages. The second — the outline design of a commercial fast breeder — was to have been completed next March, followed by a preconstruction stage that includes an application for a construction licence. The costs are being shared by National Nuclear Corporation (NNC) in Britain, Framatome in France and Siemens in Germany. Construction was supposed to begin in 1997, although that had since been put back to the end of the decade. The three governments have also funded a parallel research programme, with Britain providing a fifth of the total funding through the Department of Trade and Industry following its merger earlier this year with the Department of Energy.

Four years ago, Britain decided to terminate funding in 1994 for its own fast-reactor research programmes based on the Proto-

type Fast Reactor at Dounreay. At the time, however, it agreed to continue funding research as part of the European programme, for which it currently pays about £10 million (US\$18 million) a year.

Government officials, because of the country's economic difficulties, have now decided that even that sum is too much.



Britain's Prototype Fast Reactor at Dounreay, in Scotland, is scheduled to close in 1994.

They are planning to withdraw government funding from the European programme when the current phase ends next March — a move that will also signal Britain's withdrawal from virtually all long-term nuclear research, apart from nuclear fusion.

Energy analysts are not surprised by Britain's decision. Fast breeders are unlikely to be able to compete economically with thermal reactors until the end of the next century at the earliest, and the slowdown in the expansion of nuclear power has dimmed the prospect of a uranium shortage and the consequent need for fast reactors capable of 'breeding' their own fuel. Any decision to abandon the EFR would be welcomed by the anti-nuclear movement, which has long criticized fast reactors for their role in encouraging the spread of plutonium.

Furthermore, the privatization of Britain's electricity industry has increased efforts to find short-term solutions to the problem of energy sufficiency, and neither the government nor industry is eager to subsidize long-term commercial developments whose return remains uncertain.

Those who have been working on the fast breeder programme believe that the decision is short-sighted. They argue that, even if fast breeders are not needed until well into the next century, abandoning research at this stage would be a costly mistake.

"The NNC would be extremely dismayed and disappointed if Britain were to withdraw from the EFR programme," says Deon

Ward, general manager of NNC's fast reactor business. "We have supported this out of our own money for quite a long time. We have also developed close associations with our partner companies in other fields."

So far, Germany and France have continued to maintain their public commitment to the programme and neither is likely to dismantle its research teams immediately. However, Germany last March abandoned its own national fast breeder programme, based on the SNR-300 reactor at Kalkar, following the reactor's failure to obtain an operating licence from Social Democrat local authorities. Although Germany still spends about DM45 million (US\$30 million) a year on fast breeder research, most of this is on safety studies.

Willy Marth, executive director of the management group of the EFR, said last week that the tripartite programme was likely to collapse if Britain pulls out entirely but that the collaboration could still proceed at a slower pace if Britain were somehow to remain a partner. A spokesman for Germany's research and technology ministry, the BMFT, said that Germany's position was "not fixed, although utilities are bound to be disappointed".

Jacques Rouchard, director of nuclear reactors at France's Atomic Energy Commission, said that the remaining countries will reappportion the work between them if Britain pulls out. France, which has invested heavily in its fast breeder Superphénix, is unlikely to abandon its own fast reactor programme.

But searching questions about both the economics and the safety of fast breeders have been recently raised, the first because France has a surplus supply of electricity generated by nuclear power stations, the second because of operating difficulties at Superphénix and a recent report highly critical of its safety aspects.

In Britain, intervention by the president of the Board of Trade, Michael Heseltine — a keen supporter of European technological cooperation — could overrule the wishes of DTI officials. But this was considered unlikely.

"I would be very sad if we pulled out of the EFR", says Derek Jackson, professor of engineering at the University of Manchester, who has been studying the performance of fast reactors for both the NNC and AEA Technology. "It would be a shame if the whole programme was abandoned in the interests of short-term expediency."

David Dickson & Alison Abbott

Researchers recommend US AIDS vaccine trials

Washington. A panel of scientists and AIDS activists agreed last week that several vaccines of uncertain efficacy against human immunodeficiency virus (HIV) should be tested in large human trials. The panel's decision, if acted upon, would settle a heated dispute about the fate of \$20 million that Congress has given to the US Army to test one vaccine produced by a US pharmaceutical company. Scientists and federal officials on the panel condemned the circumvention of peer review but agreed that the money should not be rejected.

The vaccine funding measure, a late amendment to this year's defence spending bill, orders the Army to spend the money on a "large-scale...clinical investigation of the GP-160 vaccine", referring to an HIV coat protein produced by MicroGeneSys Inc. of Meridian, Connecticut. The vaccine, which is intended to boost the immune system of people already infected with HIV, is being tested in several trials.

Many scientists have argued that the research community, not Congress, should decide which drugs deserve expensive large-scale testing. They have also criticized MicroGeneSys's decision to employ Russell Long, a veteran former US senator from Louisiana, as a lobbyist. At the panel meeting, Bernadine Healy, director of the National Institutes of Health (NIH), called the

case an example of "law preempting science and scientific judgement" and said that "Congress has signed an uninformed consent form for patients with AIDS".

But the gp 160 measure also stipulates that the money will go to "other AIDS research needs of the Department of Defense" if the Secretary of Defense and the directors of NIH and the Food and Drug Administration (FDA) agree within six months after the legislation is enacted that the trial should not take place. Last week's panel was convened to advise Healy on whether the scientific data about gp 160 justifies large trials.

Scientists working with gp 160 explained at the meeting that it is too early to know whether the vaccine or any of its counterparts either prevents full-blown AIDS or prolongs lives. But panel members quickly agreed that because few good AIDS treatments exist — and because many people infected with HIV are too poor to receive good medical care — gp 160 and other promising therapeutic vaccines should be tested in massive trials.

Panel members also recommended that the trials collect other data about AIDS so that the money will not have been misspent if the vaccines perform poorly. The panel's final recommendation will be presented on 2 December at a meeting of the NIH director's advisory committee.



A vial of VaxSyn, a trial vaccine for HIV-1

The way in which the money was obtained was virtually ignored after Healy's opening speech, replaced with concern that \$20 million may not cover the cover of the planned trials. Nevertheless, it is hoped that the money will help to address both the epidemic and the demands of AIDS activists. "We hate the process, but we're focusing on what to do with the money", says Healy.

Traci Watson

US proposes relaxing rules on trials of biotech crops

San Francisco. Proposed rules that would loosen regulation of US field tests of genetically engineered crops have angered environmentalists and set industry executives worrying about a patchwork of state and federal policies. Both groups expect president-elect Bill Clinton to take a close look at the issue.

The proposal, published 6 November in the *Federal Register*, would allow US biotechnology companies to avoid a lengthy permit process and to begin certain field trials of corn, cotton, tomato, potato, soybean and tobacco with a simple notification to the US Department of Agriculture (USDA) on the same day they plant. Companies and researchers would decide whether their experiment satisfied USDA guidelines on allowable genetic constructs and their characteristics. The public has 60 days to comment.

Crops not on the list also could avoid the permit process if an institutional biosafety committee or state officials, reviewing data under the guidance of the USDA, decided that federal scrutiny was unnecessary. The

final proposal omits transgenic plants grown to harvest pharmaceuticals but recommends cautious treatment for plants incorporating functionally intact genes from human or animal pathogens.

Industry and environmentalists formed a surprising coalition in support of an early version of the regulations (see *Nature* 359, 663; 1992). But those rules were heavily revised after heated, last-minute negotiations between the agriculture department and the White House Council on Competitiveness, which has lobbied for greater deregulation in all parts of the economy.

Environmentalists say that the final revision amounts to industry self-regulation. "USDA had a germ of a good idea and the Council on Competitiveness perverted it into something that leaves little protection for the environment and the consumer", says Rebecca Goldberg, a biotechnology analyst for the Environmental Defense Fund who had supported the original proposal. The lack of advance notification will eliminate the opportunity for public comment

and make it harder to prevent abuses, she adds.

Industry officials are concerned that the proposed regulations will open the door to stricter state regulations. "We need to have a sense of satisfaction both with the general public and with state regulators", says Richard Godown, president of the Industrial Biotechnology Association. "If we don't get that, there's going to be trouble for us." John Bedbrook, research director of DNA Plant Technology Inc. of New Jersey, praises the "liberalization" of the proposed regulations but says he hopes that the federal government will require companies to provide sufficient data on crops that they wish to be exempted.

Terry Medley, chief of USDA's biotechnology section, defended the revised proposals, saying that the agency was trying to make use of available expertise by shifting some of the regulatory responsibility to state and institutional bodies. "I don't think the substance of the proposal has changed."

Sally Lehrman

Clinton's science policies and personnel have yet to emerge as he prepares for presidency

Washington. How high on the pecking order will research be in a Clinton administration? It was too soon to answer that question last week as US president-elect Bill Clinton began preparing to move into the White House.

Science was not an issue during the campaign, although both Clinton and President George Bush paid lip service to the importance of research. (Ross Perot hardly addressed the topic.) Interest in the subject was dwarfed by concern about reviving the US economy, reducing the federal budget deficit and creating a more comprehensive and less expensive health care system. As a result, anyone who pretends to know what Clinton thinks about science is putting on airs and, also, is quite likely to be wrong.

Nevertheless, there is already much speculation about the next four years. Many are optimistic about the prospects for science, although there is concern that research may have a hard time competing against other worthy demands on scarce federal dollars and, more specifically, against efforts to improve US manufacturing and repair the crumbling infrastructure. At the same time, the air is filled with anticipation about a successor to D. Allan Bromley as the presi-

dent's science adviser and whether Walter Massey and Bernadine Healy will retain their jobs as directors of the National



Al Gore is ready to implement whatever science and technology policy a Clinton administration decides it wants.

Science Foundation (NSF) and the National Institutes of Health (NIH) respectively.

Looming as a wild card is vice president-

elect Al Gore. Referred to derisively as "Ozone" by the Bush campaign because of his interest in environmental matters during his 16 years in Congress, Gore is widely expected to be a one-man clearinghouse for the president on science and technology issues. But it is not clear how Gore will do that.

"The last thing we want is another layer of bureaucracy", says one campaign adviser. "Gore doesn't want to second-guess the science adviser, although he'll want to be in the room when the science adviser talks with the president."

One possible solution is that Gore might become head of what the Clinton campaign has called an Economic Security Council, a body similar to the National Security Council that would add a technological component to the traditional review of domestic issues. Such a position would give him an opportunity to apply his knowledge and the political clout to insist that agencies carry out what the president wants.

"We don't want to do what the Bush administration did with Dan Quayle and the [National] Space Council by setting up a structure separate from the White House and NASA", says the campaign aide. "They delayed policies; we want to find a way to carry out the president's program, not get in the way."

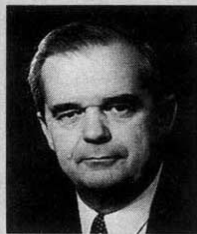
Although many names have been offered as Bromley's successor as science adviser and director of the Office of Science and Technology Policy, none has as yet caught the attention of the president-elect's transition team. Many are expecting a person from industry, in particular one experienced in manufacturing. Others believe that it is time for someone from the life sciences, perhaps biotechnology, to assume a job traditionally held by a physicist orientated towards national defence. Few expect Clinton's science adviser to be a career academic, as is Bromley.

As for the White House science office itself, Clinton's campaign rhetoric suggests that it will play an active role in coordinating technology initiatives. But reality says otherwise: the office is already too small to take on such responsibility, and Clinton has promised to reduce the federal work force by 100,000. "A bigger science office just isn't in the cards", says one aide.

What of Massey and Healy? Opinion is divided as to whether a decision to replace them would mean that Clinton has politicized the offices (Massey's term runs until 1997, and Healy serves at the pleasure of the president) or that he simply prefers fresh faces. In their favour, Healy's emphasis on the need for NIH to serve society, not sci-

Science loses two good friends

Washington. Science lost two of its staunchest advocates last week as US voters replaced more than a quarter of the 435 members of the lower house of Congress. The defeat of US Representatives Joe Early (Democrat, Massachusetts) and Bill Green (Republican, New York) seems likely to weaken congressional support for research at a vital point — within the subcommittees that set the annual budgets for the National Institutes of Health (NIH) and the National Science Foundation (NSF).



Joe Early, left, and Bill Green

throughout Massachusetts who rallied behind him "not because he delivered the goods for Massachusetts but because he understood that research was good for the country".

Green, who has served for 15 years, played a role similar to Early's as the senior Republican within the subcommittee that determines NSF's budget. Green was defeated by Carolyn Maloney, a New York City councilwoman whose efforts to link Green to his party's failed economic policies proved successful in a liberal Democratic district.

Both subcommittees will undergo wholesale changes when the new Congress convenes in January. NSF's will have a new chairman, probably US Representative Louis Stokes (Democrat, Ohio) and five of its nine members will be new.

J.M.

"It's a major blow to the biomedical research community", says Roger Fuller of the American Federation for Clinical Research about Early's defeat after nine terms by Peter Blute, a Massachusetts state legislator. "Hundreds of millions of dollars have been put into NIH and other research accounts because of his advocacy of the intrinsic value of basic research." Philip Sharp, a cancer researcher at the Massachusetts Institute of Technology, was one of many scientists

ence, is expected to be welcomed by the Clinton administration, as is NSF's current review of its mission and Massey's determination that the foundation should serve industry as well as academics.

At the same time, the two directors have very different personal styles. Healy's brashness has alienated many Democratic legislators, and there are few who would defend her if the Clinton administration decided that she is a political liability. In contrast, Massey has such a nonconfrontational style that he is unlikely to cause any problems for a president with more pressing issues on his agenda.

The future of funding for basic research is no easier to predict. Last year, in a break with tradition, Congress did not even match the president's request for NIH, and there are those who blame Healy for not fighting harder. As for NSF, the president's annual request for a double-digit increase once again was virtually ignored by a Congress that needed more money for veterans' health care, housing, the space station and other programmes that come out of the same appropriations bill.

President Bush is roundly criticized by researchers for proposing large increases for science but being indifferent to their fate. However, Republicans could always place the blame on a Congress controlled by the Democrats. Clinton will not have that option, of course, and it is expected that he will propose smaller increases but fight harder to achieve them.

Jeffrey Mervis

Fraunhofer institutes told to get more from industry

Munich. The Fraunhofer Society, which runs 47 applied research institutes in Germany, has been warned that it must look more to industry for its financing. The German Ministry of Research and Technology said last week that it will reduce its contribution to the society and suggested that the society should consolidate rather than continue to expand.

The Fraunhofer Society was founded in 1949 by industry and government as a nonprofit-making body to bridge the gap between academic institutions and industry. Its success in 'technology transfer' has been the envy of many governments, particularly Britain's, which is investigating the possibility of setting up similar systems (see *Nature* 355, 757; 1992). The number of institutes it runs has nearly doubled in the past 20 years.

The society receives its basic support from public funds, and additional money from various companies. Research minister Heinz Riesenhuber says that the society should increase the fraction of its income from contract research from 35 per cent to at least 40 per cent. Although the federal government will be making fewer grants, it will retain the

90:10 split between federal and local governments.

State funding of the institutes reached a peak this year of DM343 million (US\$220 million), including the establishment of eight institutes in the former East Germany. The government's Ministry of Research and Technology, whose own budget has been reduced, says that Fraunhofer institutes will bear a larger proportion of the cuts than other institutes conducting basic research but that plans for development of the eight new institutes will go forward.

Germany's economy is driven by small and medium-size companies that find it cost-effective to contract out their research. But the Fraunhofer Society believes that the recession will make it harder for its institutes to make up the shortfall created by government, and that some of its 7,600 employees may have to be laid off. Those companies can also look to research being done in the United States or Japan, warns the society's president, Max Syrbe, if the German government does not maintain the excellence of its applied institutes.

Alison Abbott

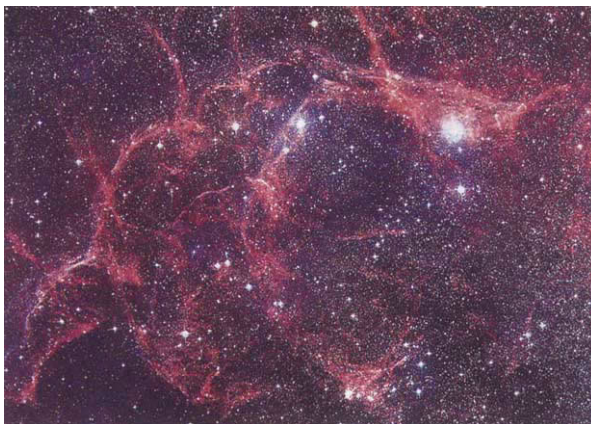
Pulsars to star in forthcoming UK music festival

London. The Lovell radiotelescope at Britain's Nuffield Radio Astronomy Laboratories at Jodrell Bank will share top billing in a concert this month. The occasion is a performance of "Le Noir del'Etoile" (The Dark Side of the Star) by the French composer Gerard Grisey at the Huddersfield Contemporary Music Festival.

The multimedia work incorporates text by French astronomer Jean-Pierre Luminet, an intricate light show projected onto an acoustic canopy that envelops performers and audience, music played by the six members of Les Percussions de Strasbourg and the sounds of two pulsars. One pulsar, in the Vela nebula in the southern skies, is prerecorded while the second, designated PSR0329-54, will be relayed by Jodrell Bank direct to the Huddersfield Sports Centre. The concert will take place on 22 November.

At two points in the performance the six musicians will fall silent, leaving the way clear for the pulsar sounds to be conveyed

over a battery of loudspeakers. Elsewhere, the radio pulsations are heard only by the musicians and provide the basic tempo for the work.



The Vela nebula: a source of sound

Grisey was introduced to pulsar sounds in 1985 by cosmologist Joseph Silk while teaching composition at the University of California at Berkeley. He likens the sound

to African drums and says he was inspired by the varying tone colour of the pulses (from the stochastic processes that generate the radio signals) and the regularity of the rhythm.

The work has been performed at Brussels and Strasbourg using live sounds transmitted by the radiotelescope at Nancy. But the superior sensitivity of the Lovell telescope and the technical lessons learnt from those performances is expected to make for a truer rendition of the artist's intentions.

The cultural excursion will not interfere with the telescope's scientific programme, according to Andrew Lyne of Jodrell Bank. Radioastronomers routinely turn the telescope dish towards various pulsars to check their timing and to look for changes in the interstellar medium traversed by the radio signals; the concert is simply a matter of squaring this particular pointing with the performance.

Roland Pease

New head of NEC research institute in US hopes to change our view of computers

Princeton, New Jersey. A determined effort by a Japanese company to create a world-class US laboratory for basic research on computers took an important step forward last week with the appointment of C. William Gear as director of the NEC Research Institute.



NEC's Princeton institute is healthy toddler.

The NEC Corporation announced on 2 November that Gear, a 57-year-old computer scientist, would lead its youngest research laboratory, formed in 1989 in Princeton. Gear joined the institute as a vice president in its first year after spending more than 25 years with the computer science department at the University of Illinois at Urbana-Champaign. With his predecessor, the late Dawon Kahng, Gear laid the foundation for a laboratory expected to generate revolutionary ideas by the end of the century about how people build, think about and use computers.

Gear says that the NEC project contains an element of enlightened self-interest along with a streak of altruism. "Just because we don't have to worry about applications now doesn't mean that we can ignore the [parent] company altogether", he says.

Even so, the differences between Gear's task and that of many other research directors can be measured in dollars and deadlines. The past year has been a tough one for research. Spending by US corporations is flat, with annual average growth slowing from 7.5 per cent between 1980 and 1985 to a mere 0.4 per cent between 1985 and 1991, according to the National Science Foundation. Japanese companies are similarly cutting back on research as their profits tumble (see *Nature* 356, 93; 1992).

But the NEC Research Institute, which officially opened its doors in May 1989, has so far been spared such cuts. Although the pace of hiring has slowed, the current staff of about 40 is still expected to grow by a third within five years. "[Hiring] nine new people a year was too fast," he says about the

original plan to reach full capacity in three years. "Five [a year] is plenty fast enough."

NEC has been generous with its US offspring, providing it with more than \$25 million last year. Although that may be only a tenth of a percentage point of its gross revenue, it is still a significant amount to spend when most companies are retrenching. And while it is dwarfed by the half-billion dollars that IBM spends annually on basic research, it is almost a quarter of what Xerox spent last year and not a bad showing for a corporate toddler.

At the same time, the cost is minuscule alongside the company's more traditional applied research and development programmes. In Japan, NEC has about a dozen such groups, which last year consumed some 12 per cent, or approximately US\$3 billion, of NEC's sales.

But that money comes with strings attached. For instance, while ideas may percolate from the research staff up to management, managers evaluate those proposals within a framework of well-defined corporate needs. Ideas that do not fit into the prescribed goals are ignored.

What is more, as NEC's profits have skittered downhill during the past year, the budgets of the applied research laboratories in Japan have suffered. Stay small, one Japanese research director advised his Princeton colleagues, and you may avoid the pain of future cutbacks.

Gear believes that stable funding is essential because scientists can become too cautious if money becomes a primary concern. He uses the carrot of steady funding to coax researchers to apply insights from one field to the unsolved problems of another.

Towards that end, the laboratory's budget is divided among individual researchers, not projects. Theorists and experimentalists are ranked according to seniority and are awarded corresponding budgets to spend as they see fit without having to provide elaborate justifications. For instance, a senior theorist may receive \$50,000 a year in addition to his or her salary, and an experimentalist more. A governing board of five to eight institute fellows approves additional expenditures for equipment.

So far, researchers at the institute have few complaints. "We've been allowed to follow our noses", says senior research scientist Eric B. Baum. Their work is spread across the spectrum of computing and physi-

cal sciences, with several scientists exploring how machines and people learn. For example, Baum is studying games such as chess and go to explain consciousness and find an alternative to the classic approach of designing a tree with branching options, which he says does not capture the way in which a go player plots strategy.

Elsewhere at the institute, scientists study the brain by asking a simple question: how does a fly keep to a straight course? They hope that the answer will help to explain the way people process external stimuli. Other researchers are exploring such problems as connecting very large numbers of processors with optical links and using computing linguistics to add more natural language interfaces into such systems.

NEC's Princeton institute is not the only US laboratory of its kind run by a Japanese corporation. Last year, Laszlo A. Belady, formerly a vice president at the Microelectronics and Computer Technology Corporation (MCC), began assembling a basic-research laboratory in Cambridge, Massachusetts, for Mitsubishi



C. William Gear

Electric. "I was practically given a free ticket," he says. His goal is to explore ways in which information technology can improve "multicultural and multinational" collaboration as well as to provide learning experiences in the work place.

Unlike Gear, Belady expects his laboratory to devote at least a third of its time to applied problems. "When they say, 'Don't worry about applications,' I say, 'Yes, sir!' but I don't believe them", says Belady. "It's not necessarily the top executives who will change their minds, but the middle management will eventually be unhappy and kill you."

Although it is not clear whether the NEC institute can maintain its emphasis on the long term if corporate profits continue to shrink, Gear says that his bosses have not changed their tune. "I'm not concerned about identifying some product" that owes its existence to work done at his laboratory, he says. However, he hopes that within five years his researchers will have changed the way NEC scientists in Japan think about the future of computing and communications.

Elizabeth Corcoran

Recession forces Japanese industry to cut back on support of universities

Tokyo. Just as Japan's cash-starved universities were anticipating extra money from the government (see *Nature* 360, 8; 1992), another source that has helped them to weather bad times over the past decade seems about to become a victim of the recession.

The 2 November issue of *Nikkei High Tech Report*, an industry newsletter, reports that companies are beginning to freeze donations to universities and to reduce joint and commissioned research with university researchers. Over the past ten years, with the encouragement of the government, donations from Japanese companies have risen steadily and now stand at nearly ¥50 billion (\$416 million) a year (see figure). That is almost as much as universities receive for competitive research grants from the Ministry of Education, Science and Culture.

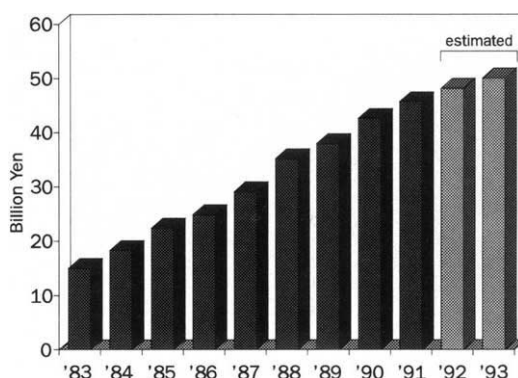
There has been a similar surge by industry in supporting joint and commissioned research with universities, although the total is only about a quarter of spending on donations. Furthermore, during the past six years the education ministry has set up 28 joint research centres on university campuses with lavish facilities to encourage collaborative research with local industry (see *Nature* 348, 7; 1990).

The donations, which constitute by far the largest contribution from industry, consist of vast numbers of awards ranging from a few hundred thousand to a few million yen (about \$1,000 to \$10,000) to individual professors. Most of the money goes to researchers in engineering faculties rather than those involved in basic research. Although the donations are small, the money does not come with the same red tape that accompanies government grants. And a few prominent professors are supported by dozens of companies.

Nikkei High Tech Report found that Tokyo Institute of Technology, Japan's leading

national university for industry-related research, received ¥2.16 billion in donations in 1991, about 5 per cent of all donations to Japan's 98 national universities. That figure rose by 20 to 36 per cent each year between 1988 and 1991. However, in the first seven

More — but for how long?



months of this fiscal year, which began on 1 April, the institute received only ¥1.16 billion, and there will be little if any growth for the entire fiscal year if donations continue to come in at the same rate.

An education ministry official disputes the *Nikkei* report, saying that generalizations cannot be made on the basis of the figures for one institution and that many companies make donations only in the latter half of the year. But he says that the ministry does not expect the growth in the donation budget to be as rapid as in recent years.

A researcher at one of Japan's leading electronics manufacturers, on the other hand, says the newsletter report is "absolutely correct". His company is reducing its donations because the sums are "so huge", he says, and because the company's research budget is being squeezed by the recession. It

is now almost impossible to recommend new funding for a university research group, he says, regardless of its quality.

The situation is not uniform across industries. Last week, 17 electronics companies reported the effect of reduced consumer demand in the first six months of the year, with Fujitsu's profits down more than 20 per cent and those of Mitsubishi Electric and NEC by nearly 30 per cent. Many electronics companies have frozen operating budgets for research and slashed budgets for buying research equipment, particularly in areas that are furthest from developing products.

In other industries, such as the pharmaceutical industry, the situation is brighter. "People have to continue buying drugs even in a recession", says one researcher from Takeda pharmaceutical company. And a professor involved in marine biotechnology research at one of Japan's national universities says that he has seen no reduction in the generous donations he gets from various industries.

The education ministry remains optimistic, opening five more joint research centres at universities and predicting a modest increase in donations from industry in 1993.

David Swinbanks

NSF commission affirms importance of basic research

Washington. The US National Science Foundation (NSF) should continue to focus on basic academic research while also supporting projects that could benefit society directly, according to members of the Commission on the Future of the NSF who attended the group's final meeting last weekend. The 15-member commission also agreed with a proposal written by member John Armstrong of IBM that the NSF should consult more often with those in the commercial sector before deciding how to spend its research dollars.

Congress has asked NSF to demonstrate its commitment to converting basic research into products in an explanation due next month of its current budget, and some of the commission's recommendations are expected to be incorporated into that document. The commission expects to deliver its report on 20 November to the National Science Board, the NSF's governing body.

Traci Watson

How not to brave the cold

Tokyo. The squeeze on research budgets in Japan's electronics companies (see above) has led to some bizarre cost-cutting measures that have backfired.

Earlier this year, Toshiba Corporation switched off the heating during the lunch hour to reduce the company's electricity bills. But chilly Toshiba researchers brought in small electric heaters and the company's electricity bills soared.

Another company tried to save money by stopping the issue of business cards to its researchers. But employees, concerned that without a name card they would cease to exist, diligently photocopied their remaining cards. The resulting increase in copying and paper costs more than offset any savings.

D.S.

French blood contamination

SIR — I should like to respond to your leading article and to the News story by Declan Butler (*Nature* 359, 759 & 764; 1992) about human immunodeficiency virus (HIV) contamination of French haemophiliacs by blood products.

I should first like to clarify my responsibilities at the CNTS (the national blood transfusion centre): I was not a "government official", I was, as a medical scientist, the head of research and development for purification of plasma proteins, not "head of research". This position placed me in the third tier of the organization. Being in research and development I had no part in collection, manufacturing or distribution of CNTS products. I cannot understand how it can be claimed that I "sold or allowed to be sold" or "supplied" any product.

In December 1984, two critical pieces of information became known to the 29 members of the AIDS-Haemophilia French study group that I coordinated and these facts were subsequently communicated (January 1985) to my senior colleagues and heads of the French Haemophilia Association. These were:

(1) French Factor VIII and Factor IX concentrates were potentially contaminated by HIV as 35–40 per cent of haemophiliacs exclusively treated with these products had antibody to LAV as it was then termed¹. Results were obtained with a research screening assay that had not at that stage been subject to rigorous analysis.

(2) In a group of 18 patients receiving the Travenol Factor VIII preparation — which was heat treated — seroconversion was not observed².

The conclusion drawn from these two observations was that the still uninfected haemophiliacs were likely to be protected from HIV infection if treated with heat-inactivated concentrates. In early 1985 there were no French heated products and only imported concentrates were available. As part of my letter of 16 January 1985 I therefore emphasized that the CNTS had responsibility in the prevention of this fatal illness (AIDS) *vis-à-vis* haemophiliacs, their doctors and the Ministry of Health. I also drew attention to the probable consequences of not signing the contract for heat-treatment technology transfer including the need for huge imports of foreign heat-treated products.

The demand for heat-treated products was also expressed by me and six other haemophilia experts in February 1985 and again by me in April following the first AIDS conference in Atlanta. These occasions are clearly documented and additional verbal but undocumented attempts on my part were also ignored.

An important difficulty in implementing preventive measures was that, until March 1985, HIV antibody assays were still under development and, until July 1985, the Abbott test was not licensed in France.

I as much as anybody at that time was deeply concerned with the risks the haemophilia population were potentially incurring. However, it is important to recognize that by the earliest date at which heat-treated products would have been introduced, that is, if the advice of myself and others had been implemented forthwith, more than 90 per cent of those who were later established to be infected had in fact already seroconverted — that is, had already been infected.

In brief, the difficulty I face today is to get people to understand that my conclusions in early 1985 were not perceived then as self-evident facts beyond scientific criticism — hence the difficulty in getting my advice implemented. Uncertainties existed at that time that have now been resolved — uncertainties about the accuracy of serological tests, about their clinical significance, about the risk of seropositive patients developing AIDS and the efficacy of various heat-treatment processes. The information available at the time was shared by a large number of scientists, medical specialists and officials at the French Haemophilia Association through their medical advisers. The problem everywhere, not just in France, was to convert information into prompt action in organizations that were traditionally bureaucratic.

J-P. Allain

University of Cambridge,
Department of Haematology,
MRC Centre, Hills Road,
Cambridge, CB2 2QH, UK

1. AIDS-Hemophilia French Study Group. *Blood* 66, 891–901 (1985).

2. Rouzioux, C. et al. *Lancet* i, 271 (1985).

Cave of the ostrich

SIR — Irrespective of whether George William Stow produced a spurious Bushman painting of a man disguised as an ostrich approaching a flock of real ostriches, there is no doubt that he visited a cave which he designated 'the cave of the ostrich'. He listed it in his list of important caves of the Bushmen¹. He also mentioned that it was near Oliphant's Been.

Oliphant's Been is, in fact, a pillar of rock standing some 15 metres high with weathered rocks and rockface around it, close to the border of Lesotho in the

southern Orange Free State. I visited the site in December 1982 and found the cave within about 10 metres of the rock pillar. I doubt whether many people had been there during the 100-year interval between Stow's visit and mine. Apart from mentioning the site in his list of notable caves, Stow also mentioned it briefly as being near an 'excellent ochreous' earth supply². As it was an extraordinarily hot day, I did not look for the ochre deposit after photographing the paintings but regard this mention as definite evidence that Stow visited the site.

As far as I am aware, these paintings have never been reproduced and, in fact, Stow's title should be in the plural as there are actually three ostriches in the cave and not much else. Two of the ostriches are painted together and the other, with an extremely attenuated neck, is separate. There are one or two rather crude red paintings, one of which might also represent an ostrich, and a very faded human figure, but the two ostriches in black with touches of white on the wings, obviously males, are the main attraction of the site.

H. C. Woodhouse

1 Buckingham Avenue,
Craighall Park,
Johannesburg, 2166 South Africa

1. Stow, G. W. *Native Races of South Africa* (Struik, Cape Town, 1905, reprinted 1964).

2. Stow, G. W. *Report of the Geological Surveyor* (OFS Newspaper Company, Bloemfontein, 1878).

Misquoted

SIR — Sir Hermann Bondi (*Nature* 358, 363; 1992) noted that "Thus science is not something strange and odd but the most human of pursuits. Hoyle, in his novel *The Black Cloud* puts this so well when a nonscientist says of scientists: 'I cannot understand what makes them tick. They are always wrong and they always go on.' This very popperian sentiment inspires us all through our trials and tribulations. Popper has made it clear that we should be proud so to be described."

Remarkably, there is no such statement in *The Black Cloud*. The nearest thing to it in the edition published in 1957 (Harper and Row, New York) is on page 166, where Francis Parkinson, the private secretary to the Prime Minister, said "It isn't so much the volume of talk that surprises me" he laughed. 'We get plenty of that in politics. It's the number of mistakes they've made, how often things have turned out differently to what they've expected.'

Robert E. Davies

Department of Animal Biology,
University of Pennsylvania,
Philadelphia, Pennsylvania 19104, USA

The case for conservation

SIR — There is stronger case for conservation than that recently argued by Henry Gee (*Nature* 357, 639; 1992). Gee rightly stresses thinking in terms of systems rather than lists and that “the integrity of the system as a whole transcends the extinction of particular species”. But it is for this very reason that it is hard to justify large investments in preserving particular endangered species. In fact a system may not only be preserved but also improved when particular parts are replaced by more efficient ones. This may be what is happening in our own ecosystem as traditional extinction through climatic change and evolution through random mutation and natural selection are being replaced by animal and crop husbandry and genetic engineering.

These trends suggest that Gee may be wrong to suggest that there is a “vested interest of human populations in keeping things very much as they are”. In fact, new varieties, such as thriftier cattle and more disease-resistant grains, seem more in our interest than the traditional varieties. And there seem to be some species, such as the human immunodeficiency virus, whose extinction would be much to our advantage. So intelligent, responsible human intervention in evolution seems more in order than is mere conservation.

The problem, however, is that human intervention in evolution may tend to produce varieties that will be increasingly unable to fend for themselves and increasingly dependent upon husbandry. Valerius Geist (*Nature* 357, 274–276; 1992), for example, has recently pointed out that hybrids of white-tailed and mule deer “suffer from severe deficiencies when it comes to tactics and strategies of predator avoidance”.

The more the particular parts of the global system are replaced by varieties that depend on husbandry for their survival, the more the system as a whole may depend upon the intelligence and responsibility of the human species. But humans cannot always be trusted to “have dominion over the Earth” (Genesis, I, 26) in an intelligent and responsible manner. Greed, stupidity, wars and so on often render us unwilling or unable to husband nature for the long-term good either of the system as a whole or of the human species.

Because our intelligence and goodwill are erratic at best, we need the safety mechanism of vast numbers of traditional species and varieties, which do not depend upon husbandry, in order to increase the chances of the whole sys-

tem's surviving even when we humans are behaving at our worst. This is the objective case for conservation.

Frank J. Leavitt

Jakobovits Center

*for Jewish Medical Ethics,
Faculty of Health Sciences,
Ben Gurion University of the Negev,
PO Box 653,
84105 Beer-Sheva, Israel*

Anyons

SIR — John Maddox (“Time for tempting Nobel fates” *Nature* 359, 101; 1992) misspells anyons for ‘anions’.

The semantic content of this term, most frequently associated with fractional statistics and anyonic superconductivity, may therefore have been lost on the uninitiated reader.

Stanley Wu-Wei Liu

*American Physical Society,
1 Research Road,
PO Box 1000,
Ridge, New York, USA*

Too many reviews?

SIR — Have we become review crazy? Suppose one is interested in learning or writing about axogenesis and plasticity in the developing visual system. Using Melvyl, one scans first for keywords in titles back to 1987 and then asks for reviews among those articles. Under plasticity, there are 1,860 articles, 373 of which are reviews, under vision 7,261 articles of which 653 are reviews; synapse 1,000 to 126; axon 3,012 to 192; neurotransmitter 3,642 to 632; LTP 634 to 75; retina 7,416 to 489; neuronal development 2,277 to 340; extracellular matrix 4,764 to 749; neurotrophic 565 to 73; neurobiology 370 to 175; laminin 2,385 to 157. This averages out to nearly 12 reviews for every 100 articles published, or one review for every 8 articles published.

Three aspects of this trouble me. First, in writing a review, political acumen dictates the inclusion of most papers, regardless of quality, even papers for which retractions have been offered. This uniform treatment tends to reduce the impact of quality work and elevate marginal or outdated work. Second, reviews tend to replace the actual articles they review and when referenced lend the impression that the author of the review and not the researchers themselves is responsible for the work. Finally, this whole trend tends to trivialize science into the equivalent of journalistic ‘sound bites’ — they are becoming the Cliff notes or cribs of the overworked scientist seeking a short cut.

Is this the best response to the overwhelming surge in papers published and

the growth of interdisciplinary fields? Have randomly designated researchers become spokespersons for fields by writing reviews, and do those reviews then become the sole source of information? At one time a yearly issue of *Annual Reviews* in a couple of fields would suffice, but now at least a dozen journals have arisen solely devoted to a constant, monthly issuance of reviews (for example *Current Opinions*).

Is quality science and a critical reading of the original literature best achieved through a rear-view mirror? Maybe more thought should be given in the future to both the writing and solicitation of reviews.

Michael J. Ignatius

*Department of Molecular and Cell Biology,
University of California, Berkeley,
Berkeley,
California 94720, USA*

The welfare state

SIR — Barry Hughes¹ credits me with an aphorism usually attributed to H. L. Mencken. He then proceeds to give simple answers to complex questions. He lists only three variables, including ambient temperature, to which egg production is sensitive (true) and he says these do not have “much impact on welfare”. But freezing and excessively high temperatures are both deleterious to animal welfare and to comfort. Other variables that lower egg production include infectious disease, sudden fright, harassment by roosters and parasitism. Hughes says that I argued that “if production reaches some target value which is commercially satisfactory, it demonstrates that welfare also is satisfactory”. But I made no such argument: I said that *impaired* production resulted from adversity²; and that “production values have been an important adjunct to aiding animal welfare”³.

Hughes cites his own review of experiments that “demonstrate that egg output of caged hens declines progressively as the number of hens per cage is increased and also as the space per hen is reduced”. This supports my statements that egg output is reduced by discomfort, and that “alleviation [of unsatisfactory caging for hens] should be obtainable by improved husbandry”. My ‘bottom line’ was “Farm animals should be treated humanely. Their well-being is necessary for productivity”².

Thomas H. Jukes

*Department of Integrative Biology,
University of California, Berkeley,
Berkeley,
California 94720, USA*

1. Hughes, B. O. *Nature* 359, 472 (1992).

2. Jukes, T. H. *Nature* 355, 582 (1992).

3. Jukes, T. H. *Nature* 358, 186 (1992).

Prospects for women in science

SIR — Your recent leading article on women in science¹ reflects insensitivity to the problems women face in science, in marked contrast to recent leading articles in *Science*² and in the *Lancet*³. The assumption in your leading article and in a recent report from the Max Planck Society⁴ is that the only problems stopping women making it to the top in science are child care and long working hours. This one-sided emphasis is an intellectual insult to those women — with or without children — who are already working more than 40 hours a week and are striving to make an independent career in science. Such women share with the writer of the leading article the belief that “there is no evidence that sex is related to success in scientific research” and are prepared to be judged by the same objective standards as their male colleagues. However, in return, women have the right to demand the same job opportunities and the same resources, and to enjoy the same privileges as are given to men at similar stages in their careers. Women in science can face prejudice that may be “entirely unintentional” but is nevertheless real, and should not be comforted with the statement that “there is no doubt that fewer women than men are among those at the top of their profession and that this situation should change with the changing numbers of women in the workplace”. Of course it should but will it?

For instance, only 2 of 210 directors (or heads of research groups) in the Max Planck Society (MPG) are female⁴ (although it should be noted that the only female director in the natural sciences has won the Lasker award). One cannot apply for a vacant position as an MPG director, and although the system used to select new directors has obvious advantages in choosing fields and selecting able individuals, a figure of 1 per cent suggests that women are not properly considered. In addition, the trend in the MPG to establish theme institutes rather than select the best individuals regardless of field further discriminates against women, as women who might be considered as directors are not distributed equally over all fields.

In 1991 the MPG used positive action to establish 29 five-year groups in the former East Germany and all are headed by men. Women are also poorly represented in the German scientific academies — the Bavarian Academy of

Sciences, for example, had in 1991 not only no female member but had had none for at least the past ten years. Should one really accept the implicit but never stated argument that in all these instances not a single woman can be found without lowering the current standards? Is such an argument tenable in a profession that otherwise prides itself on objectivity of judgement and international standards? Or does it merely reflect a desire to preserve the status quo?

A serious effort to improve the situation in Germany or elsewhere has to approach the problem at two levels. On the one hand ways have to be found to identify and appoint qualified women to top positions, and to set and publicize realistic short-term goals (*not* quotas) to increase the number of women at the top who participate in decision-making processes and who act also as role models. As shown by experience in the United States, this can be done without a decrease in standards. On the other hand, positive action should be used for a limited time at a lower level to increase the pool of qualified female scientists from which future holders of top positions — be it in research institutes, in universities, or in industry — can emerge. One way might be to finance starting scientist positions that are preferentially given to women and which could be held for five years in a laboratory of the individual's choice. This would give more women sufficient resources, time and self-confidence to develop their own research programmes at a critical point⁵ in their careers. The Habilitation Program of the Deutsche Forschungsgemeinschaft, which is preferentially aimed at women and which also includes some funds for child care, is a step in the right direction. Other measures that force consideration of whether there are qualified women — whether for speakers at a meeting as in the case of the National Science Foundation, for a tenured appointment as at good universities in the United States for the past 20 years, or on committees for evaluation of institutes or departments — can only be welcomed as they draw attention to an imbalance that is at best a long way from being solved⁶, and at worst, at least in Europe, is too often swept under the carpet and ignored.

Mary Osborn
Max Planck Institute
for Biophysical Chemistry,
W34 Goettingen, Germany

Letters submitted for Correspondence should be typed, double-spaced, on the side of the paper only.

1. *Nature* **359**, 92 (1992).
2. *Science* **255**, 1333 (1992).
3. *Lancet* **337**, 1007–1008 (1991).
4. *MPG Spiegel* **2** (91), 18–20 (1991).
5. Smith, D. O. *Science* **257**, 1597 (1992).
6. Mason, J. *Nature* **353**, 205–206 (1991).

Altered food

SIR — It is unfortunate to encounter evidence of more silliness about new biotechnology applied to food (“Chefs forswear genetically altered food,” *Nature* **359**, 8; 1992). Perhaps as scientists we take the prospect of “new biotech” or “genetically engineered” foods for granted because they are so obviously refinements of the older ones that have been with us for so long. Yeast has been used to brew beer for eight millennia, and farmers were crossbreeding livestock long before Mendel’s experiments helped to define the principles of genetics. For decades, genes have been transferred from one species to another and even from one genus to another to yield commonly available food plants including oats, rice, currants, potatoes, tomatoes, wheat and corn (Goodman *et al.* *Science* **236**, 48; 1987). These plants, resulting from wide crosses made possibly by embryo rescue techniques, are “genetically engineered” and even “transgenic” by any reasonable scientific definition. They are found not just in laboratories or test plots but are the very same fruits, vegetables and grains sold at the local supermarket. The recombinant DNA and related techniques of the “new biotechnology” essentially speed up and increase the precision of gene transfer.

Far from eliciting concern, the use of more precise techniques that yield a better-characterized and more predictable plant variety should be seen as potentially positive. It was therefore surprising to see a group of chefs, of all people, resisting further improved varieties. Undoubtedly, much of the nonscientists’ uncertainty about biotechnology results from a lack of perspective on its pedigree; would they boycott the genetic hybrid of tangerine and grapefruit that we call a “tangelo”? Or the mutant peaches we call “nectarines”?

If a few sadly exploited chefs really do intend to make do with “non-genetically engineered” foods, their menu is likely to be a singular one, offering little more than fish, shellfish and wild game and berries. I’ll order from the *Menu de Degustation de Genetechnologie*, thank you very much.

Henry I. Miller
Office of Biotechnology,
Food and Drug Administration,
Rockville, Maryland 20867, USA

SIR — Jeremy Rifkin’s *Foundation on Economic Trends* is shortsighted in its campaign against genetically engineered foods.

I plan to boycott restaurants that display the Pure Foods Campaign logo.

Joseph J. Brigham
3103 Coriander Drive,
New Bern, North Carolina 28562, USA

A polarity in European research

Lennart Philipson

The large Human Capital and Mobility programme may at first sight seem like good news for European scientists. But this is far from the case.

THE European Commission has recently launched a major fellowship programme, called Human Capital and Mobility, that could easily interfere with and even jeopardize other, similar European programmes directed by European scientific organizations. The total amount of money available for the Human Capital and Mobility programme is around 340 million ECU (\$428 million), which compares, for example, with around 6 million ECU per year spent by the European Molecular Biology Organization (EMBO) for its fellowships. One consequence of this vast capital investment is that the stipends from the European Commission are much higher than those from organizations like EMBO: the amount paid for an European Commission postdoctoral stipend is more comparable to a professor's salary than to that of a postdoctoral fellow.

The European Commission has made assurances that the scientific standards of this programme should be high and rigidly adhered to during the assessment of applications, but the application and the peer-review procedures would surely be considered a parody by anyone outside the system. A major fraction of the funds in the first round was reserved for fellowships to so-called centres of excellence, yet these applications had to be submitted within two weeks of receipt of the application forms. The space available on these forms allowed only a summary presentation of the project, and no specific questions were asked about the selection procedure for the fellows which, of course, has to be very stringent to ensure the highest quality.

Inconsistencies

The peer-review committee for the Commission's fellowships, consisting of scientists of high reputation, was not allowed to evaluate the proposals independently, but its members were asked to come to Brussels for several days to read and comment on the proposals. Several of these scientists could not stay for more than a single day and the final allotment of funds seems to indicate that the committee has drizzled support over as many centres as possible instead of identifying a few top ones. The European Commission is obviously again confusing two different intentions: what it needs to do is to develop a system for awarding separate types of support to developing

and élite science centres.

As things stand, after passing this initial scrutiny, each selected centre had to rewrite its application in a contract form, now taking into account the bureaucratic rules of the Commission that were not mentioned in the original application forms. These rules may make it impossible for some centres to accept the money they have been awarded.

Next, and most important, this massive influx of fellowship support may very well lead to reluctance on the part of some European governments to provide funds for fellowships to European organizations like EMBO or the Human Frontier Science Program. Even the small amount of money provided for research through other, smaller European organizations may be jeopardized when the resources are compared with those provided by the Commission. It is understandable that European governments, which pay through different channels for all these European activities, feel that fellowship funds in Europe must soon become saturated because of this massive influx from the Commission. It would be a tragedy if support for these existing European programmes begins to fall away, because the selection of fellows through a reliable peer-review system and the flexibility in providing funds is much better in several of these smaller programmes.

Although additional resources for scientific training and mobility in Europe are always welcome, this latest example of the European Commission's idiosyncrasy appears to establish, after numerous similar incidents in the past, that correction of the system can hardly be achieved from within the organization. It obviously does not realize it is sick.

Representing an organization independent of the Commission, I propose an alternative mechanism to assure support both for the Commission and its competing European institutions before Europe becomes locked in a centralized, fossilized policy. Each of the current 12 member states of the European Commission should pay a contribution to the Commission consisting of at least 1% of collected value-added tax and in addition 90% of the import duty on goods imported from outside the European Communities. All this may in the future amount to up to 1% of the gross national product of the individual member states.

This major contribution must, of course, be based on a well-controlled income declaration which must be negotiated between Brussels and the governments of the member states.

Survival

To ensure that alternative European programmes survive and receive adequate support despite these gigantic transfers to Brussels, a simple solution would be that the Council of Ministers which, after all, if unanimous, is the body determining the Commission's policy, requests that all the costs for joint European activities should be deductible from the income-tax return to Brussels in the same way as an individual taxpayer may deduct expenses incurred for earning the income. With this system, the European Commission's budget would probably be affected very little. This system would also certainly improve national resources for basic science within member states. The European programmes would furthermore be better coordinated with the Commission's activities, since the same national agency would probably be in charge of both the payment to Brussels as well as the deduction from this payment. It would probably take only one meeting at the prime-ministerial level to ensure that science programmes at least operate with a decentralized, functional system instead of a bureaucratic, inflexible European Commission policy.

Therefore, with a single move one can establish two important principles: first, that European initiatives are not competing with each other for funds but are allowed to develop in a decentralized fashion; and second, that these different activities are compared with regard to quality, cost-effectiveness and administrative efficiency as they would be evaluated by the same agencies at national levels. After all, Europeans can afford to test out several different procedures to achieve first-class science. It is high time that we reorganize European science to regain the leading role that Europe has had in the past. □

Lennart Philipson is Director-General of the European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, Germany. The opinion expressed in this commentary is personal and does not reflect the official position of EMBO or EMBL.

Towards real-time molecular demolition?

A demonstration of the successive excitations of rubidium atoms by lasers may be a model for the selective dissociation of molecules for which chemical engineers would (and will) give fortunes.

EVEN as we marvel at how lasers are being used for the manipulation of atoms or molecules, we are assaulted by further illustrations of the tricks the laser people are waiting to play on us. Now, of course, it is a commonplace that people working with atomic or molecular beams can arrange, by irradiation with laser light of the appropriate frequency, that much of their material is in a predetermined excited state. Because the energy levels of atoms or molecules depend to some small degree on the masses of their nuclei, the same principle may yet be used commercially for the separation of isotopes. And those whose expertise is the use of insubstantial electromagnetic fields for trapping small numbers of atoms or molecules in ionic form can now use lasers detuned from some known resonant frequency to 'cool' their specimens even to the microkelvin level.

All that, and much more, is familiar, but there is plainly no end to what lies in store. Here are some recent innovations, each a surprise in itself and also suggestive of other tricks not far ahead.

Take, for example, the dissociation of molecules, which could be of great importance in the chemical industry if it could be used to push chemical reactions in some desired direction. In principle, there is no problem. Identify an excited state of the molecule in which its components are no longer bound together, work out the frequency corresponding to the excitation energy and then irradiate the material to be dissociated with laser light of that (or greater) frequency.

But there are snags. The frequencies are usually far in the ultraviolet (requiring the inconvenient use of excimer lasers, themselves dependent on chemical reactions), the process is inefficient and the products of dissociation often unspecific. So why not, instead, pump laser energy into particular vibrational states of molecules, literally to shake them apart in such a way that the products are specifically determined by the particular chemical bond attacked in this way?

On the face of things, there should be no difficulty. For one thing, molecular vibrations usually have frequencies in the infrared. Then, to a first approximation, a molecular vibration is a simple harmonic process, with the consequence (described in every textbook of quantum mechanics) that the energy levels are equally spaced. But reality differs from

the first approximation. By definition, if a molecule can shake itself apart, its potential energy cannot be a symmetrical function of the departure (for a diatomic molecule) of the internuclear distance from its equilibrium value. For molecular vibrations in general, the energy difference between successive excited states decreases as the energy increases.

So why not arrange that molecules meant to be shaken apart should be irradiated not by laser light of constant frequency but with light of a frequency that decreases as time goes by? Two years ago, S. Chelkowski, A. D. Bandrauk and P. B. Corkum suggested just that (*Phys. Rev. Lett.* **65**, 2355; 1990). Their idea was to use femtosecond laser pulses with a pre-patterned declining frequency to shake molecules apart. It is easily foretold that some future generation of chemical engineers will really be laser people in disguise, skilled at shaking molecules apart in specific ways by laser pulses of varying frequency. (The technical term is 'chirping', borrowed from acoustic analysis of bird song.)

A Dutch group has now carried through a rudimentary demonstration of the principle of chirped laser excitation, not with molecules, but with rubidium atoms. The point is simply that the valence electron of rubidium, whose ground state is the $5s$ state, may be excited to $5p$ and then $5d$ by optical photons, with wavelengths of 780.2 nm and 775.9 nm respectively. The first transition is less energetic than the second by a little more than 0.5 per cent.

The difference is so small that a single laser might span the gap, with a little manipulation of its output. B. Broers, H. B. van Linden van den Heuvell and L. D. Noordam from the Institute for Atomic and Molecular Physics at Amsterdam have now shown how that can be done (*Phys. Rev. Lett.* **69**, 2063; 5 October 1992). Chirping, it appears, works as predicted. But there is a big surprise.

The source of the laser light is a pulsed dye laser, made to yield four different outputs by means of dye cells pumped by an yttrium-aluminium-garnet laser. The four sequences of pulses are mixed together in a water cell so as to produce a relatively broadband laser pulse lasting a few tens of times as long as the original (depending on the geometry), and with a predominant frequency increasing or

even (according to the geometry) decreasing over the duration of the output pulse. That is a hint of the technology the next generation of chemical engineers will have to practise.

The Amsterdam group has amply demonstrated what its readers will have expected of it: with an upward (infrequency) chirping pulse, it turns out to be possible to excite rubidium atoms with great efficiency into the second excited state ($5d$). So much can be inferred from the arrangements for monitoring the population of that state, which consist of the ionization of excited rubidium atoms by more energetic laser light (at a wavelength of 532 nm) and the counting of the electrons therefrom. Upward (in frequency) chirped laser pulses behave as expected.

So where is the surprise? Chirped laser pulses whose frequency decreases over the duration of the pulse also seem to have the same effect, but even more quickly. (Remember, the first rubidium transition is less energetic, by about 0.5 per cent, than the second.) The authors describe this observation as "counter-intuitive". That is like applying the same adjective to an observation that water will spontaneously run uphill.

That the counter-intuitive mechanism is nevertheless plausible has been shown by calculation. Low-to-high frequency chirping seems fully to populate the intermediate state. Chirping in the other direction can fully populate the upper state without foot touching ground at the intermediate level. In reality, of course, none of this would be possible unless the laser pulses spanned a range of frequencies throughout their short duration.

The implications of these developments for the next generation of chemical engineers are not easily discerned. For most of us, it will be a comfort that the idea of frequency chirping as a systematic way of dissociating atomic systems has been substantiated. But the application of the technique now demonstrated to real chemical systems will evidently require a much more detailed knowledge of the details of the structure of complicated molecules than even future decades will easily acquire. Perhaps the solution is to use laser pulses shorter than a femtosecond in duration to shape the femtosecond laser pulses that shake molecules apart. Real-time molecular demolition may be the name.

John Maddox

Stitch in time saves design

S. J. Judge

THERE is currently considerable enthusiasm for 'neural network' models of brain function, but one issue that has not been resolved is how the structure of real brain circuits (which are by no means always the densely randomly interconnected networks favoured by the theoretical network modellers) can be incorporated into such models. Perhaps because its circuitry is relatively well known, the eye-movement system has attracted several attempts to marry the network approach to the particulars of real brain circuitry and function¹⁻³. The latest of these attempts is to be found on page 159 of this issue⁴, where Lisberger and Sejnowski offer an intriguing model of the vestibulo-ocular reflex (VOR).

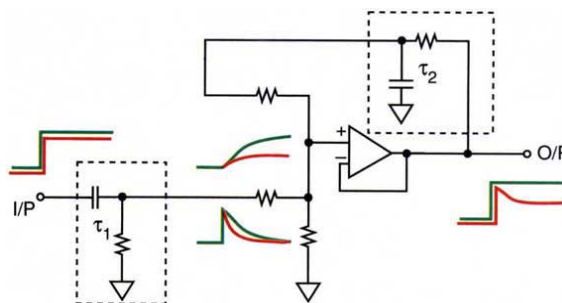
The VOR is the reflex that prevents the image of the world sliding over the retina during head turns. Head rotation is sensed by the semicircular canals in the inner ear, and that signal is used to produce eye movements that rotate each eye in its socket so as to keep eye angular velocity in space close to zero during head turns. Because even low-velocity image movement impairs visual resolution, VOR gain (eye angular speed divided by head angular speed) should be kept close to unity.

Brain circuitry

There has been a longstanding dispute about the nature of the brain circuitry that adjusts VOR gain, with Ito⁵ championing the view that the VOR gain is adjusted by regulating the size of the vestibular signal flowing from the canals through a particular part of the cerebellum called the flocculus and back to the main pathway in the brainstem. The difficulty with accepting this argument is that by far the most thorough study of the behaviour of neurons in the flocculus, by Miles and his colleagues⁶, found that although the signals there did change when the VOR gain was manipulated (by experience in various optical devices that altered the relationship between head movement and its visual consequences) the signals changed in the wrong direction for Ito's hypothesis. An alternative model was therefore proposed⁷ in which although gain does change at the connection between the incoming vestibular fibres and flocculus neurons, the key factor in altering VOR gain was regulation of the size of the vestibular signal flowing through the main brainstem pathway.

Lisberger and Sejnowski now present a very simple model which may be able to resolve the conflict between the two groups, and which also incorporates a

principle that the authors suggest may have a more general application. This model is not itself a 'neural network' model, but a classic engineering model of a kind made familiar through the influential work of Robinson⁸. Lisberger and Sejnowski propose that, as well as gain changes in both the direct brainstem and cerebellar pathways, there is a third modifiable element, namely the speed of the dynamic response in the pathway carrying vestibular input to the cerebellum. Lest one gives the impression that Occam's razor is being neglected, let it



Electrical equivalent of Lisberger and Sejnowski's proposed neural circuit for the VOR. If the time constants τ_1 of the input filter and τ_2 of the filter in the positive feedback loop are equal the circuit will have unity gain, as shown in the waveforms in green. If τ_1 is less than τ_2 the circuit will respond to a step with the waveform shown in red and will have a d.c. gain of less than unity.

be said that the model requires the two gains to be matched, or it will become unstable at low frequencies. How the system manages to keep these gains matched is not explained.

To see how the dynamic element works, consider the electrical realization of Lisberger and Sejnowski's model (see figure). The vestibular input is filtered by a high-pass filter and summed with the output of a low-pass filter, whose input is itself the output of the whole circuit. (The unity gain amplifier is present only as a buffer.) The form of the output depends on the mismatch between the time constants in the two filters. If that of the first filter is smaller, then the response to a step input will be a step plus exponential slide in the same direction. If the time constant of the first filter is larger, then the response will be a step plus exponential slide back. In other words, the circuit converts a change in the dynamic properties of the input filter to a change in steady state gain. It is important to note that this brings about gain change without the adjustment of any gain 'knob' in the circuit, and it is this curious feature of the circuit that gives it the potential for

reconciling the models of Ito and Miles *et al.* When the VOR gain is reduced, this is achieved by the combination of a modest reduction in gain in both the direct and cerebellar pathways, and a much more dramatic reduction in time constant of the cerebellar input. The latter allows the VOR gain to be reduced despite the fact that the gain of the connection between the vestibular input and the flocculus has changed in the 'wrong' direction. Why the gains should change at all is not explained.

Hidden variable

This idea is simple and ingenious, and of course one wishes it well. But it works by introducing a 'hidden variable' to account for otherwise contradictory data

— although this is always a possible resolution of a contradiction, one wants to see direct evidence for the postulated change in dynamics. Although Lisberger and Pavelko⁹ have evidence that the dynamics of the eye-movement response to a head movement do change in a way compatible with the new model when VOR gain is changed, it is by no means certain that the neural input to the flocculus does alter its dynamics in the way supposed.

The circuit is of course a very odd one to the electrical engineer because its

design feature is just the opposite of normal engineering practice, where one seeks to arrange the circuit so that its overall performance will be reasonably consistent despite incidental variations in the values of circuit components. This circuit has just the opposite property, in that it exploits the fact that small mismatches of components have drastic consequences. There is of course a price to be paid for such a design: the circuit is only quasi-stable. But, biologically, perhaps that is a price worth paying. □

S. J. Judge is in the University Laboratory of Physiology, University of Oxford, Parks Road, Oxford OX1 3PT, UK.

1. Anastasio, T. J. & Robinson, D. A. *Neural Computn* **1**, 230–241 (1989).
2. Robinson, D. A. *A. Rev. Neurosci.* **12**, 33–45 (1984).
3. Lisberger, S. G. & Sejnowski, T. J. *Univ. California San Diego Inst. Neural Comput. Tech. Rep.* 9201 (1992).
4. Lisberger, S. G. & Sejnowski, T. J. *Nature* **360**, 159–161 (1992).
5. Ito, M. *A. Rev. Neurosci.* **5**, 275–296 (1982).
6. Miles, F. A. *et al. J. Neurophysiol.* **43**, 1406–1493 (1980).
7. Miles, F. A. & Lisberger, S. G. *A. Rev. Neurosci.* **4**, 273–299 (1981).
8. Robinson, D. A. *A. Rev. Neurosci.* **4**, 463–503 (1981).
9. Lisberger, S. G. & Pavelko, T. A. *J. Neurosci.* **6**, 346–354 (1986).

Phoenix galaxies

Lars Hernquist

A COLD, massive cloud of molecular gas close to, but not part of, the prominent galaxy M81 has been discovered by the radioastronomers N. Brouillet, C. Henkel and A. Baudry¹. As the first intergalactic molecular complex to be discovered, the cloud is of interest to those cosmologists wondering how the primordial material of the Universe originally collapsed to form galaxies. But its vicinity to M81 will endear it to astrophysicists who argue that dwarf galaxies of similar mass are born phoenix-like out of the debris of galaxy–galaxy collisions.

The observable mass in galaxies like the Milky Way resides almost entirely in stars and interstellar gas, the latter comprising both atomic and molecular components. It is widely believed that stars form in dense, cold molecular complexes. By extrapolation to much larger scales and to earlier times in the history of the Universe, it is natural to suppose that the stars in other galaxies originated in molecular gas. Thus, galaxy formation may have followed a simple progression whereby protons and electrons made in the Big Bang combined to produce atomic gas, which then cooled to yield star-forming regions of molecular gas. These considerations have motivated astronomers to search for intergalactic gas, which might represent a precursor of galaxies. Atomic gas has indeed been detected in many groups and clusters of galaxies², but not molecular gas lying between galaxies, at least until the new observations of Brouillet *et al.*

Employing detectors sensitive to rotation-line emission from carbon monoxide at 115.3 and 230.5 gigahertz, Brouillet *et al.* mapped the molecular distribution in a relatively nearby group of galaxies which includes the prominent spiral galaxy M81. This collection is distinguished by the presence of faint structures apparently linking some of its members. Numerical simulations demonstrate that these ‘bridges’ and ‘tails’ are produced during interactions between galaxies³, implying that tidal effects have played an important role in the evolution of the M81 group. In the course of their observations, Brouillet *et al.* discovered a complex of molecular gas in the vicinity of M81, but not lying within its disk or nucleus. Moreover, this gas does not appear to be physically associated with any of the dwarf galaxies in the group. The authors are thus led to conjecture that this gas is in fact intergalactic in nature; if so, this represents the first actual detection of molecular gas outside the main body of a galaxy.

The molecular complex observed by

Brouillet *et al.* has a mass between 10^6 and 10^7 solar masses, which is similar to the largest agglomerations of molecular gas in the Milky Way. Although not part of M81 *per se*, the complex is nonetheless bridged to this galaxy by a faint wisp of atomic gas which extends further out towards two other members of the group, M82 and NGC3077 (see figure). The authors suggest that this arm is of

for studying how mostly gaseous objects turn into stars, especially in the context of tidal interactions. At present, there is no stellar source associated with the molecular complex, but this situation will change in relatively short order. Because the complex is spatially confined, the stars produced in it will probably make a gravitationally bound object. Eventually, it may well resemble the young dwarf galaxies in the M81 group, suggesting that we are seeing a new galaxy arising from the debris produced by the collision of M81 with one of its companions. If so, then, as the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Radio map of neutral hydrogen in and between the galaxies M81 (right) and NGC3077 (lower left). The molecular complex discovered by Brouillet *et al.* is in the knot of emission (yellow/green) just to the left of the spiral galaxy M81 and to the right of the centre of the figure. (J. M. van der Hulst and R. I. Allen/Science Photo Library.)

tidal origin and was drawn from the disk of M81 during a close passage of one of these companion galaxies. Numerical models indicate that the collision probably occurred at least 10^8 years ago, which is much longer than the lifetimes of molecular complexes in the Milky Way. This, combined with the lack of a molecular bridge to M81, suggests that the complex formed *in situ* — that is, in intergalactic space.

The prevailing conditions in the M81 group may not reflect those characteristic of gas from which galaxies like the Milky Way formed, as the gas in the molecular complex discovered by Brouillet *et al.* has already been chemically processed by a galaxy; namely M81. Nevertheless, the M81 group can rightly be thought of as a miniature laboratory

authors suggest, this molecular complex may be the missing link between the atomic gas and the young dwarf galaxies in the group.

The notion that some dwarf galaxies might be born during the deaths or transformations of more massive galaxies is an old one. Indeed, Zwicky⁴ was among the first to suggest that the diffuse tails torn from disk galaxies during collisions may include debris that evolves into objects resembling dwarf galaxies like those orbiting the Milky Way. Interest in this proposal has been renewed in recent years by the discovery of ‘condensations’ or ‘knots’ in the tails of many systems that apparently consist of two spiral galaxies in the final stages of merging. In particular, Schweizer⁵ showed that the condensation in the

interacting pair NGC4038/9 exhibits features characteristic of young stars, and Mirabel *et al.*⁶ further showed that these stars must have formed *in situ*, much like the molecular complex now discovered near M81. As argued by Mirabel *et al.*, the properties of the knot in NGC4038/9 are typical of dwarfs and they speculate that it represents the genesis of a new galaxy.

Similar structures have also been reported⁵ in the merger candidate NGC-7252 and more recently^{7,8} in the infrared-luminous system IRAS 19254-7245. In particular, the tails in the latter object possess at least nine condensations, some of which are sites of vigorous star formation⁷. Zwicky's conjecture has also been boosted by numerical simulations of galaxy mergers. At least two independent groups have now established that tails produced by tidal interactions of galaxies are subject to fragmentation, yielding structures not unlike those observed^{9,10}.

The significance of these observations and calculations for the origin of most dwarf galaxies is problematic. Violent collisions, like those in NGC4038/9, for example, completely destroy both progenitor disks, leaving remnants that are akin to present-day elliptical galaxies. Clearly, the dwarfs around the Milky Way could not have formed in this manner, as most of the stars in our Galaxy reside in a thin disk, as in other normal spirals. In addition to possibly bridging the gap between the molecular gas prerequisite to star formation and the young dwarfs in the M81 group, the observations by Brouillet *et al.* suggest that Zwicky's mechanism may operate in encounters much weaker than those in fiducial merger candidates. To the extent that future observations and numerical modelling support this viewpoint, it may well be that many dwarf galaxies were not formed primordially, but are a consequence of the violent dynamical nature of the Universe. □

Lars Hernquist is at the Lick Observatory, University of California, Santa Cruz, California 95064, USA.

A retrograde step forward

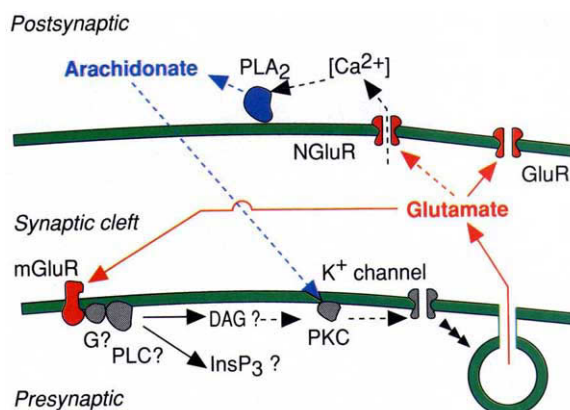
David G. Nicholls

MOST excitatory synapses in the mammalian brain use the amino acid glutamate as their neurotransmitter. Changes in the effectiveness of glutamatergic synapses are believed to underlie many of the processes of information storage (learning) and retrieval (memory), although the relative importance in this synaptic plasticity of presynaptic glutamate release as opposed to postsynaptic modification is still a matter for

arations of isolated nerve terminals.

Synaptosomes are isolated nerve terminals obtained by the careful homogenization of brain regions. They remain functional *in vitro* for several hours and retain all the machinery for the uptake, storage and release of their neurotransmitters. But because they are no longer attached to an axon they must be stimulated artificially. Herrero *et al.* used 4-aminopyridine, a nonspecific

inhibitor of fast-activating K^+ channels that induces spontaneous action potentials in synaptosomes³. Nerve terminals contain α -, β - and γ -isoforms of protein kinase C (ref. 4) which can be activated *in vitro* by phorbol esters. In the presence of 4-aminopyridine (but not of elevated KCl, the more conventional way of depolarizing such preparations), protein kinase C activation causes a burst of glutamate transmitter to be released from synaptosomes⁵. This potentiation of secretion by protein kinase C is well documented in a number of cells, including the platelet and chromaffin cell, where the kinase seems directly to potentiate the intracellular coupling between Ca^{2+} and secretion. In contrast, electrophysiological experiments with neuronal cell bodies indicate that protein kinase C may modulate the activity of ion channels. Linden *et al.* show⁶ that in the cell bodies of Purkinje cells, which express the γ -isoform of protein kinase



A speculative scheme for the role of arachidonate in presynaptic metabotropic receptor-mediated positive feedback. During normal transmission, glutamate exocytosed into the synaptic cleft activates the postsynaptic AMPA/kainate glutamate receptor (GluR) responsible for normal transmission but fails to activate the NMDA receptor (NGluR) owing to the latter's voltage-dependent block by Mg^{2+} . The presynaptic metabotropic receptor (mGluR) does not cause inhibition of the K^+ channel because protein kinase C (PKC) is relatively insensitive to G-protein activation in the absence of arachidonate. Under conditions of repetitive stimulation (dashed arrows), when NGluR is activated, Ca^{2+} entry through this receptor would activate postsynaptic phospholipase A_2 (PLA₂), so generating arachidonate which may then diffuse to the presynaptic membrane, sensitizing PKC and coupling the mGluR to protein kinase C (PKC) by the action of phospholipase C (PLC) and diacylglycerol (DAG). The resulting inhibition of the delayed rectifier would potentiate presynaptic action potentials and further enhance the release of transmitter. Inositol trisphosphate (InsP₃) would not participate in this model.

lively debate. A problem has been the lack of any plausible mechanism for regulating the release of glutamate in response to an action potential. Arachidonic acid is one candidate that has been put forward as a 'retrograde messenger': this would be produced on the postsynaptic side of the synapse and then diffuse back across to the presynaptic terminal to enhance the release of glutamate¹. Now, on page 163 of this issue², Herrero *et al.* describe their discovery of a positive-feedback mechanism that operates in the presence of low concentrations of arachidonic acid and potentiates glutamate release from prep-

C, the delayed rectifier K^+ current is potentially inhibited by phorbol ester. Unsaturated fatty acids can also activate the kinase and inhibit the channel⁶. Protein kinase C in the nerve terminal functions by a similar mechanism, apparently potentiating the action potentials induced by 4-aminopyridine by inhibiting a K^+ channel that controls repolarization⁵.

The contribution of Herrero *et al.* is, first, to identify the presynaptic receptor responsible for the activation of protein kinase C and, second, to show that the coupling of receptor to channel is dependent on low concentrations of arachidonic acid. The phospholipase C-

- Brouillet, N., Henkel, C. & Baudry, A. *Astr. Astrophys.* **262**, L5-L8 (1992).
- Silk, J., Djorgovski, S., Wyse, R. F. G. & Bruzual, G. A. *Astrophys. J.* **307**, 415-425 (1986).
- Toomre, A. & Toomre, J. *Astrophys. J.* **178**, 623-666 (1972).
- Zwicky, F. *Ergebnisse der Exakten Naturwissenschaften* **29**, 344-385 (1956).
- Schweizer, F. in *Structure and Properties of Nearby Galaxies* (eds Berkhuysen, E. M. & Wielebinski, R.) 279-285 (Reidel, Dordrecht, 1978).
- Mirabel, I. F., Dottori, H. & Lutz, D. *Astr. Astrophys.* **256**, L19-L22 (1992).
- Mirabel, I. F., Lutz, D. & Maza, J. *Astr. Astrophys.* **243**, 367-372 (1991).
- Colina, L., Lipari, S. & Macchetto, F. *Astrophys. J.* **379**, 113-121 (1991).
- Barnes, J. E. & Hernquist, L. *Nature* (in the press).
- Elmegreen, B. G., Kaufman, M. & Thomasson, M. *Astrophys. J.* (in the press).

David Bohm (1917–1992)

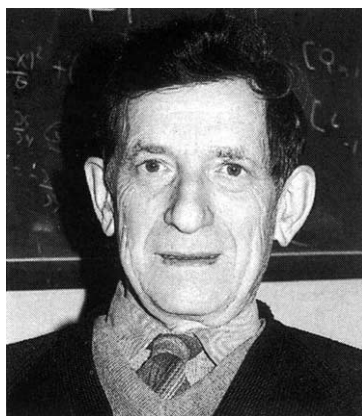
THE death of David Bohm on 27 October is a sad loss to the community of physicists, and particularly to those concerned with fundamental issues in quantum mechanics.

In straight physics, Bohm's seminal contribution was made in conjunction with his student Yakir Aharonov: the discovery in 1959 of what is now known as the Bohm–Aharonov effect. This predicted the effect of a flux of magnetic induction confined to the interior of a long solenoid on the relative phases of electron waves passing either side of the solenoid, in regions of space where the magnetic induction was zero. This was the first demonstration of how a nontrivial topological feature of the vacuum (in this case identified with the absence of magnetic induction) could produce striking physical effects. Taken together with earlier work on plasma oscillations and the behaviour of electrons in metals, this assured Bohm's reputation as a master of insightful physics in the mainstream tradition.

But in the longer term his exhaustive investigations on unorthodox interpretations of quantum mechanics may be seen as an even more significant contribution. Dissatisfied with the orthodox Copenhagen interpretation, which he had beautifully expounded in his text book *Quantum Theory* in 1951, he published a preliminary investigation of what came to be known as the causal interpretation in 1952. This was a hidden-variables approach, in which the exact location of the particle played the role of the hidden variable and the

standard quantum-mechanical predictions were deduced by including an extra term, the 'quantum potential' in the classical Hamilton–Jacobi equation. Effectively, the theory involved waves and particles instead of the orthodox waves or particles (an idea which actually goes back to de Broglie in the 1920s).

The remarkable feature of Bohm's interpretation was that it apparently circumvented the famous proof given by



von Neumann in 1932 that hidden-variable interpretations of quantum mechanics were mathematically impossible. The situation here was finally clarified in the 1960s by the late John Bell, who showed that von Neumann's proof involved overly strong and physically unmotivated assumptions, which Bohm's theory naturally denied, but the price to be paid for the otherwise appealing hidden-variables approach

was a feature of instantaneous nonlocal action mediated by the quantum potential. In a deep sense this violates the spirit of relativity theory, although Bohm was able to show that at the level of statistical observation, all the predictions of relativistic quantum mechanics could be accommodated.

What was the reaction to the causal interpretation? On the one hand Bohm was attacked for being reactionary, of trying to understand quantum mechanics as a sort of glorified statistical mechanics; but on the other hand he was also attacked for his more extreme philosophical views which he had developed in parallel with his causal interpretation. These views distinguished an explicate order, as he called it, of apparently independent entities, from an underlying implicate order involving a view of ultimate reality in which everything was bound up holistically with everything else. His critics regarded this, quite unfairly, as akin to oriental mysticism. However, during the past few years the causal interpretation has been gaining increasing 'respectability' with physicists, not least perhaps because it was enthusiastically championed by Bell, a great admirer of David Bohm.

At the time of his death Bohm was in the process of completing, with his main collaborator Basil Hiley, a definitive presentation of his approach, and this, when hopefully it is published, will be the monument by which Bohm's ideas must ultimately be judged. That I am sure is how he would have wanted it himself.

M. L. G. Redhead

coupled metabotropic glutamate receptor⁷ acts through a G protein to activate phospholipase C and generate the two second messengers inositol trisphosphate and diacylglycerol (which in turn activates protein kinase C). Herrero *et al.* show that agonists for the receptor mimic the action of phorbol ester in enhancing the 4-aminopyridine-induced release of glutamate, although the intervening stages in the protein kinase C activation are not yet defined.

The presence of a presynaptic autoreceptor (one that responds to the transmitter released by the terminal) that stimulates rather than inhibits release is unusual. Most transmitters are subject to a negative-feedback control by autoreceptors that prevent further release by activating K⁺ channels and/or inhibiting Ca²⁺ channels. But such a mechanism acting at the glutamatergic terminal would effectively prevent the terminal from responding to a high-frequency stimulus — the condition necessary for the induction of long-term potentiation (LTP), the most investigated form of synaptic plasticity⁸. On the other hand, a

permanently engaged positive-feedback regulator could lead to the glutamatergic synapse running out of control. It is here that the second aspect of the work of Herrero *et al.* becomes relevant. They show that the metabotropic glutamate receptor seems to be coupled to K⁺-channel inhibition only in the presence of arachidonic acid. Arachidonic acid can be produced in neurons by phospholipase A₂ activated by Ca²⁺ entering through the NMDA (N-methyl-D-aspartate)-selective glutamate receptor⁹. The idea that arachidonic acid acts as a retrograde messenger, whereby it is generated postsynaptically and then diffuses across the synapse to increase glutamate release¹, is based on the integral part played by the NMDA receptor in the establishment of LTP. But until now we have had no convincing explanation of how this messenger function might be executed.

Although Herrero and colleagues are not directly concerned here with synaptic plasticity, their report does provide groundwork on which constructive speculation can be based (see figure). Re-

gardless of whether the ultimate expression of plasticity is presynaptic or postsynaptic, if synaptic plasticity is all-or-none at a given synapse then it is important that once a synapse becomes committed to potentiation, for example by NMDA-receptor activation, the release of glutamate should remain maximal for the period needed to establish plasticity. Arachidonate could assist this process by recoupling the positive-feedback pathway found in the presynaptic terminal. □

David G. Nicholls is in the Department of Biochemistry, University of Dundee, Dundee DD1 4HN, UK.

1. Williams, J. H. *et al.* *Nature* **341**, 739–742 (1989).
2. Herrero, I., Miras-Portugal, M. T. & Sánchez-Prieto, J. *Nature* **360**, 163–166 (1992).
3. Tibbs, G. R. *et al.* *J. Neurochem.* **53**, 1693–1699 (1989).
4. Shearman, M. S., Shinomura, T., Oda, T. & Nishizuka, Y. *J. Neurochem.* **56**, 1255–1262 (1991).
5. Barrie, A. P., Nicholls, D. G., Sánchez-Prieto, J. & Sihra, T. S. *J. Neurochem.* **57**, 1398–1404 (1991).
6. Linden, D. J. *et al.* *J. Neurosci.* **12**, 3601 (1992).
7. Sladeczek, F. *et al.* *Nature* **317**, 717–719 (1985).
8. Collingridge, G. L. & Singer, W. *Trends Pharmac. Sci.* **11**, 290–296 (1990).
9. Lazarewicz, J. W., Wroblewski, J. T. & Costa, E. *J. Neurochem.* **55**, 1875–1881 (1990).

Molecular genetics arrives on the farm

James E. Womack

THE appearance of transgenic mice on the covers of *Nature* and *Science* in 1983 prompted predictions of an immediate revolution in animal agriculture. As a result of molecular genetics and biotechnology, it was said, super cows, pigs and other species carrying transgenes to make them unusually productive and healthy would dot the pastoral landscape by 1993. But although transgenic mice have revolutionized the study of mammalian development and gene expression, the technology has never made it to the farm.

Fortunately, the potential for molecular genetics in animal improvement was perceived by some to be far broader than that provided by the trial-and-error protocols of gene transfer. Now, the fruits of such vision are finally being realized — in papers published last month, Shuster *et al.*¹ described the identification of the genetic defect that causes leukocyte adhesion deficiency (LAD) in Holstein cattle, and Rudolph *et al.*² the identification of the mutation responsible for hyperkalaemic periodic paralysis (HYPP) in Quarter Horses (so-called because they were originally bred to race over a quarter of a mile). Together with earlier work on porcine malignant hyperthermia^{3,4}, these studies stand as landmarks in comparative medicine.

The two groups^{1,2} began with the observation of animal phenotypes clinically similar to diseases described in humans (mutations responsible for human LAD and HYPP had already been identified in loci coding CD18 β -integrins and the sodium channel α -

subunit, respectively). Both groups sought and found polymorphism in the animal homologues of these loci, established genetic linkage of the disease with these polymorphisms in segregating families, and then sequenced genes from normal and diseased animals to establish the molecular basis of the animal disease. Bovine LAD was shown to result from an aspartic acid to glycine substitution at amino acid 128 in CD18, a region where several mutations have been found to cause human LAD. The sodium-channel mutation in equine HYPP substitutes leucine for phenylalanine in a transmembrane domain thought to lie near the cytoplasmic surface. Two human mutations causing HYPP are also expressed near the cytoplasmic surface, although in a different domain of the protein. Sequencing of the normal and mutant genes in cattle and horses has resulted in diagnostic methods, based on the polymerase chain reaction, that can unambiguously identify carrier animals.

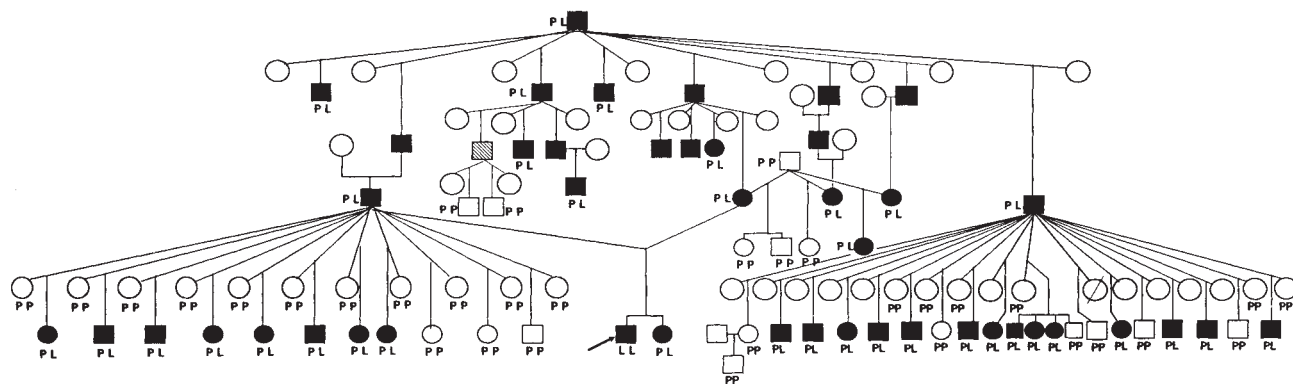
The two papers also reinforce the importance of comparative medicine in general and comparative genetics in particular — clinical human and veterinary medicine provided a similarity of disease phenotypes and, in this case, human medical genetics provided the candidate genes for the animal diseases. The upshot is not only diagnostic tools that may be worth millions of dollars (bovine LAD costs the dairy industry about \$5 million a year in the United States alone) but also two animal models defined as mutations in the homologues of mutant human genes, an important

finding for the extension of experimental medicine to the level of gene therapy.

Animal breeding practices commonly increase the genetic relatedness of individuals within a breed because of the extensive use of select sires. The carrier frequency for the mutation in bovine LAD, for example, is 15 per cent among bulls and 6 per cent among cows, all related to one outstanding bull. Equine HYPP was traced to a common sire which has contributed to the genotypes of 2.9 million horses. These breeding practices explain how an allele can permeate a breed in relatively few generations. They also offer encouragement for rapidly replacing an unwanted allele with the help of proper tools for carrier identification.

Somewhat disconcerting, however, is the unresolved issue of possible physiological associations of the disease genes and the phenotypes responsible for the rapid global spread of germ plasm. Pronounced muscularity is a highly desirable phenotype in Quarter Horses. The sodium-channel mutation responsible for HYPP may redirect ion traffic in such a way as to induce hypertrophy and thus create an inseparable linkage of the disease and a valued phenotype. The potential beneficial effects of the carrier state of the mutant CD18 gene in cattle are not obvious, but there is precedent for heritable association of disease and production in the Weaver syndrome in Brown Swiss cattle⁵. The questions that must be answered are whether these positive/negative associations exist and, if they do, whether they are physiologically intertwined or merely genetically linked and thus separable by recombination.

These two success stories emphatically validate the candidate-gene approach to resolving the mutational basis for inherited animal diseases and selectively removing them from breeding populations.



The HYPP Quarter Horse pedigree of Rudolph *et al.* (reproduced from ref. 2). Horses were typed for the normal allele of the sodium channel (phenylalanine, P) and the mutant allele thought to cause HYPP (leucine, L). Heterozygous horses are afflicted with HYPP (black circles and squares), as is a horse homozygous for leucine (arrowed); the disease status of the male denoted by the hatched square is unknown. Squares, males; circles, females.

Equally valid is the genomic approach to animal improvement, which is also rapidly coming of age in that linkage maps of highly polymorphic markers are being developed in a number of agriculturally important animals. Most economic traits, including both positive and negative traits controlled by one or more genes, are not represented by obvious candidate loci at the moment. A saturated map of markers applied to families segregating economically important traits will produce linkage associations of traits to markers, resulting initially in the use of the marker for selective breeding and perhaps eventually in the positional cloning of the genes involved. In work in press and as yet unpublished, our group has collaborated with Georges and others to map the Weaver syndrome⁶ and the polled (hornless) gene in cattle to microsatellite markers, now permitting marker-assisted selection on the basis of approximately 3 per cent and 13 per cent meiotic recombination, respectively. Many more marker-trait associations in cattle, pigs and chickens can be expected in the very near future as marker maps of these species emerge and are applied to families segregating interesting traits.

Despite the parallels between the bovine LAD and equine HYPP stories, the authors of the papers present a markedly different outlook for dealing with the two diseases. Shuster *et al.* point to the potential for eradicating bovine LAD within one year. Rudolph *et al.* mention the same possibility for equine HYPP, but temper it with reservations as to how the horse industry will use the tool. Both are probably test cases for the effect that the advent of molecular genetic markers will have on the respective animal industries.

We may or may not see transgenic livestock in our pens and pastures in 2003. The gate is open, however, for the elimination of inherited diseases and marker-assisted selection of loci responsible for economically valuable traits. Molecular genetics has arrived on the farm. □

James E. Womack is in the Center for Animal Genetics, Institute of Biosciences and Technology and the Department of Veterinary Pathobiology, Texas A&M University, College Station, Texas 77843, USA.

Ignition of X-ray flares

Matthew G. Baring

PUZZLING, rapid variations of intensity of X-rays coming from active galactic nuclei are explained in this issue in terms of unstable interactions between photons and relativistic protons originating in the galaxies' cores. Writing on page 135, Kirk and Mastichiadis¹ show that this mechanism can account for the onset of X-ray outbursts. Because they also predict the X-ray spectrum to be expected from the mechanism, their model has the advantage of being testable. Also, because the X-ray characteristics will depend on the galactic nuclear environment, the variability may prove to be a probe of active galactic nuclei.

The X-ray emission from active galactic nuclei is thought to emanate from volumes barely greater than the Solar System (10^{10} km across) around a supermassive black hole at the hub of each galaxy. Such compact sources are inevitably involved if the rapid variations in intensity are to be explained. Learning how the X-rays are produced will ultimately reveal how the black holes interact with their surrounding medium, ionizing it and driving the remarkable phenomena of active galactic nuclei.

The model of Kirk and Mastichiadis, which builds on work of the past decade, is in particular aimed at explaining the most extreme variability, X-ray flaring for durations as short as an hour, seen in active galactic nuclei. It describes an instability by which a population of relativistic protons interacts with radiation to generate a rapid escalation in the radiation intensity; shortly afterwards the instability will saturate, or burn itself out, and the plasma will return to its more stable conditions.

The relativistic protons are possibly produced by so-called Fermi acceleration at plasma shock waves² created in the dynamic environs of the central black hole; this acceleration process is commonly invoked to explain the generation of cosmic rays. These protons then interact with 'seed' ultraviolet photons, which are expected to be found in an accretion disk around the black hole, to create relativistic pairs of electrons and positrons (this occurs by the electromagnetic process $\gamma p \rightarrow p e^+ e^-$, where γ is a photon, p , a proton and $e^\pm$ are the electron and positron). The pairs then spiral through the magnetic field and lose energy by radiating X-rays (as synchrotron radiation). The instability can set in because the X-rays can then collide with the protons to produce further electron-positron pairs, which radiate more X-rays, and so on, quickly draining energy from the protons and putting it

into the X-ray flux. The rate of X-ray production is proportional to the product of the X-ray density and the density of relativistic protons so that, if the latter exceeds some critical value, the whole proceeds catastrophically to create a flare. This is exactly what Kirk and Mastichiadis find, with a critical proton density of 10^8 – 10^9 cm⁻³ for central regions of active galactic nuclei. The ultraviolet photons do not participate in the feedback loop, and so act merely as seeds for the instability.

This description of the instability is somewhat simplified. One complication is that the pair-production process $\gamma p \rightarrow p e^+ e^-$ has a threshold, and will proceed only if the product of the photon and proton energies exceeds the square of the electron rest-mass energy. For X-ray photons this implies that some of the protons must have energies of a few teraelectronvolts or more, a mild constraint for protons accelerated by shocks. Furthermore, if the X-rays are synchrotron photons, then their energy depends on the magnetic field strength, so that there is also a weak dependence of the minimum proton energy on the field strength.

A key feature of the model is that it predicts an X-ray spectral index (logarithmic slope of the photon energy-flux distribution) that is relatively insensitive to the steepness of the accelerated-proton distribution, and is generally in the range 0.5–1.0 observed in active galactic nuclei. At the same time, the critical proton density corresponds to a maximum energy density of protons that is permitted in the quiescent pre-flare phase, which is correlated with the pre-flare X-ray compactness (or energy density). This characteristic should be a strong indicator of which active nuclei should exhibit flaring activity. For those that do, deductions about the proton distribution can be made using the pre-flare phase observations.

The work of Kirk and Mastichiadis deals specifically with the linear first stage of the instability. A natural extension is to follow the evolution of the photon and proton populations through to saturation. If the proton-acceleration mechanism occurs at a fairly constant rate, then the proton density gradually builds up to the critical value and then the instability triggers. As the X-ray flare grows, the proton cooling rate must increase and eventually exceed the acceleration rate, thereby draining the relativistic-proton population and eventually terminating the instability. Time-dependent models should make valuable

1. Shuster, D. E., Kehrli, M. E. Jr, Ackermann, M. R. & Gilbert, R. O. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9225–9229 (1992).
2. Rudolph, J. A. *et al. Nature Genet.* **2**, 144–147 (1992).
3. Fujii, J. *et al. Science* **253**, 448–451 (1991).
4. Otsu, K., Khanna, V. K., Archibald, A. L. & MacLennan, D. H. *Genomics* **11**, 744–750 (1991).
5. Hoeschel, I. & Meinert, T. R. *J. Dairy Sci.* **73**, 2503–2515 (1990).
6. Georges, M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).

predictions of the time-evolution and duration of flares and their correlation with spectral index and pre-flare X-ray compactness, suitable for direct comparison with observations.

Other complications need to be addressed to refine the flare model. The most successful theory of the quiescent phases of active galactic X-ray activity is focused on steady-state electromagnetic pair cascades^{3,4}, where photons and pairs interact through Compton scattering (electron-photon collisions), and pair-producing photon-photon collisions. Such cascades efficiently generate X-rays and rapidly cool pairs and protons, and could significantly hamper both the pre-flare accumulation of protons (and consequently affect the triggering of the instability), and also alter the evolution of the flare. The inclusion of photons below ultraviolet energies to model radio-loud sources could also be very influential.

The extension of Kirk and Mastichiadis's model to include time dependence is already in progress, as is the develop-

ment of more complete time-dependent models including pair cascades^{5,6} that will yield more powerful diagnostic information from comparisons between theory and observation. Interesting by-products of these models include neutrons and neutrinos that are spawned by pions created in photon-proton interactions^{1,6}. Modelling X-ray variability with these developments promises exciting times ahead for astrophysicists who are probing the interior regions of active galactic nuclei. □

Matthew G. Baring is in the Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA.

1. Kirk, J. G. & Mastichiadis, A. *Nature* **360**, 135–137 (1992).
2. Kazanas, D. & Ellison, D. C. *Astrophys. J.* **304**, 178–187 (1986).
3. Lightman, A. P. & Zdziarski, A. A. *Astrophys. J.* **319**, 643–661 (1987).
4. Zdziarski, A. A. *et al.* *Astrophys. J.* **363**, L1–L4 (1990).
5. Stern, B. & Svensson, R. in *Relativistic Hadrons in Cosmic Compact Objects* (eds Zdziarski, A. A. & Sikora, M.) 41–50 (Springer, Berlin, 1990).
6. Stern, B., Sikora, M. & Svensson, R. in *Testing the AGN Paradigm* (eds Holt, S. S., Neff, S. G. & Urry, C. M.) 313–316 (Am. Inst. Phys., New York, 1992).

ECOLOGY

A leaf for all seasons

Peter D. Moore

THE form and physiology of leaves vary according to the environment within which they develop — leaves used to sun or shade, together with humid or arid conditions, display a whole range of morphological and biochemical adaptations to different light intensities and moisture regimes. In environments where there is strong seasonal variation in these conditions the type of leaf formed may vary with the season, especially as a response to drought. One habitat where one might not expect such adaptation to occur is the humid understorey of tropical forest in Central America, but it seems that it does. Mulkey and colleagues, writing in *Proceedings of the National Academy of Sciences*¹, describe their observations on the shrub *Psychotria marginata* (Rubiaceae), and their finding that it produces drought-resistant leaves in anticipation of the annual dry season.

Leaves are extremely plastic organs. Even within a single plant they can often vary in their structure according to their position, especially on large, long-lived perennials such as trees. The complexity of architectural structure in a tree leads to the generation of its own heterogeneous microclimate and the conditions experienced by leaves in the lower layers of branches differ in many respects from those of the upper canopy. Light intensity is lower, humidity is

generally higher and less variable, temperatures are lower and more uniform, wind speeds are lower, and so on.

Trees often respond to these variations in microclimate by producing leaves with different morphological and biochemical properties at different positions in their canopies. Those leaves that originate in the shade tend to be less thick, and have a thinner epidermis, fewer stomata and a lower density of hairs, all characters associated with the enhancement of light trapping and a low emphasis on water conservation. Owing to the high degree of structural complexity, these features of leaf variation within individual trees are particularly apparent in tropical rain forests². Such intra-plant variation in leaf form may relate to the genetic mosaics³ known to exist among many perennial plants, but work in this area is still in its infancy.

Drought resistance in leaves can also vary within a single plant temporally as well as spatially. In regions with strong seasonal climate contrasts, such as areas with Mediterranean climates — hot dry summers and cool moist winters — a range of leaf strategies is available, including the deciduous and evergreen options⁴. The production of different types of leaves with morphological and physiological features appropriate for the contrasting seasons has been recorded from such habitats⁵.

Harper⁶ has pointed out that the leaves of trees can be considered as populations and has shown that leaves are produced in cohorts that subsequently age and die together, a feature that is only too obvious in deciduous trees, but is less clearly evident in evergreens. Phases of leaf birth need not be an annual event but may be more frequent, as Mulkey and colleagues¹ have now observed in the tropical evergreen shrub *P. marginata*. Their work was carried out in Panama, and it turns out that in the *P. marginata* understorey leaves are produced in two waves, the first in May/June and the second in December/January (at the beginning and the end of the wet season respectively). From January through to April, the shrub experiences a degree of lower water availability during the 'dry' season; precipitation for the 13-week dry season is about 88 mm, compared with an annual total of 2,600 mm, but humidity is always in excess of 65 per cent. Demographic studies of these two cohorts of leaves show that they survive for at least two years. Photosynthetic assimilation rates are the same in both sets of leaves, but those leaves produced immediately before the dry season have lower transpiration rates and higher water-use efficiencies while maintaining their photosynthetic production. This differential is maintained even in plants that are irrigated throughout the dry season.

There are obvious advantages for a tropical forest plant in producing a cohort of leaves capable of conserving water during such a dry spell, especially in years of the Southern Oscillation when the dry season may be extended. The maturation of fruits also corresponds with the beginning of the dry period, and extra water is then required for the production of these fleshy organs. 'Wet season' leaves are generally produced in greater quantities (up to 70 per cent of total leaf production), but there are evident benefits in maintaining the two leaf types as a hedge against excessive drought. The advantages of retaining wet-season leaves are more subtle and may be related to the higher stomatal conductance of this leaf type during periods of high water availability, leading to higher carbon dioxide concentrations in the mesophyll.

The production of 'dry season' leaves does not seem to be related to the onset

1. Mulkey, S. S., Smith, A. P., Wright, S. J., Machado, J. L. & Dudley, R. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9084–9088 (1992).
2. Roth, L. *Stratification of Tropical Forests as Seen in Leaf Structure* (Junk, The Hague, 1984).
3. Gill, D. E. in *Plant Ecology* (ed. Crawley, M. J.) 321–343 (Blackwell Scientific, Oxford, 1986).
4. Moore, P. D. *Nature* **312**, 703 (1984).
5. Ehleringer, J. *Am. J. Bot.* **69**, 670–675 (1982).
6. Harper, J. L. in *Plant Canopies, their Growth Form and Function* (eds Russell, G., Marshall, B. & Jarvis, P. G.) 105–128 (Cambridge University Press, 1989).
7. Wright, S. J. *Ecology* **72**, 1643–1657 (1991).

of drought, for it precedes this event and is maintained even under irrigated conditions (though prolonged irrigation for several years does interfere with the synchrony of leaf production⁷). There appears, therefore, to be an endogenous rhythm associated with the generation of the two types of leaf. The occurrence of this heterophylly in a rain forest plant that is subject to only very mild seasonal

drought is in itself intriguing, but it also raises questions about palaeoclimate. How long, for instance, has this region been subject to a dry season? That in turn begs the question of how long it takes for such heterophylly to evolve. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

felt by people living on the coast, in contrast to other events of a similar size along the Mid-American trench. The weak shaking is attributed to smooth slow rupture propagation, which may be related to slip extending through low-velocity material all the way to the sea floor. Shallow slip is an efficient generator of tsunamis. However, some of the largest tsunamis are believed to be related to slumping of soft sediments near the trench. As the Nicaraguan trench has little sediment, the tsunami there illustrates that even zones without soft sediments can produce large tsunamis if the faulting extends to the sea floor (Kanamori).

This variability poses a problem for the recently contested seismic-gap hypothesis (see Seth Stein's News and Views article³), which states that sites for future earthquakes are likely to be found in regions, termed seismic gaps, along a major plate boundary that have not recently ruptured in major earthquakes (S. P. Nishenko, US Geological Survey, Golden, Colorado). However,

EARTHQUAKES

The exception is the rule

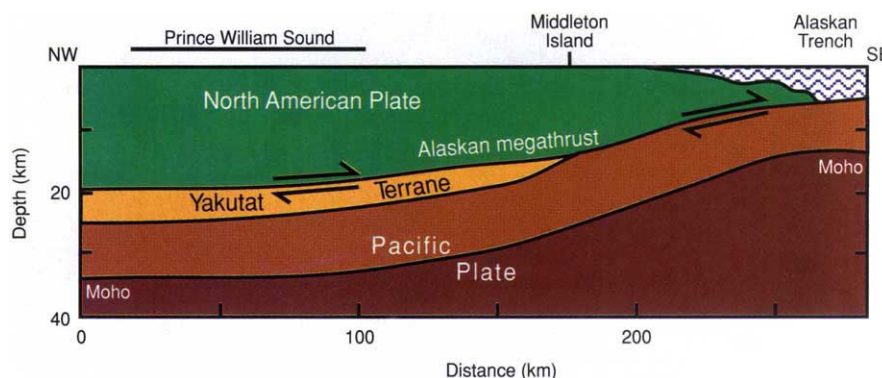
Heidi Houston

GREAT earthquakes on subduction zones, such as those along the Pacific rim, can cause death and damage not only through the direct shaking but also by the tsunami waves they can generate. Accurate forecasts could reduce such losses, but seismological crystal balls are clouded by two problems explored at a recent conference^{*}: how can the locked and dangerous segments of a subduction zone be distinguished from those slipping smoothly and safely; and how can accurate forecasts be made when these locked segments may rupture differently from cycle to cycle? Efforts to address these problems show greatest promise at subduction zones for which modern seismicity can be compared with other data such as slip distributions in past earthquakes, plate-boundary structure, pre-historic seismicity and modern strain accumulation.

A decade of advances in determining details of the rupture of great earthquakes (those whose magnitude exceeds 8.0) has resulted in an increased appreciation of the complexity of this process. In the few places where the historical record spans several cycles of great earthquakes, the variability in the mode of rupture is striking. Several factors, in addition to the inter-event time, can vary from cycle to cycle, including whether a region ruptures in one event or more, the direction and velocity of rupture, the distribution of slip, and whether the shallow part of the subduction zone slips (H. Kanamori, California Institute of Technology; S. Beck, University of Arizona). These factors can significantly affect seismic and tsunami hazards.

For example, in 1906 a 500-km segment of the Ecuador-Colombia subduction zone ruptured in a magnitude 8.8 earthquake. The same segment slipped in a sequence of events in 1942 (magnitude 7.6), 1958 (7.7) and 1979 (8.2), for which the seismic moment released totalled only about one-fifth of the 1906 event (Beck; ref. 1). Because the earlier

earthquake ruptured a longer region at once, it was also able to rupture over a greater width down-dip and with a greater displacement. Along the Nankai trough in southwest Japan, historical data spanning 1,300 years with 11 great and several large earthquakes reveal two occasions on which the entire 500-km segment ruptured at once, and four



Inferred crustal structure in the Prince William Sound region of the Alaskan subduction zone from TACT seismic sounding. A continuous, gently-dipping reflector near the depth of the 1964 Alaska earthquake epicentre is interpreted as the megathrust (major interplate boundary and slip plane of the 1964 earthquake) separating the North American plate from the subducted Yakutat terrane, a microplate that has become entrained between the North American and Pacific plates and appears sutured to the Pacific plate in this region⁴. Currently seismicity is concentrated in the subducting package and exhibits down-dip tension; the megathrust itself is seismically quiet, accumulating strain. (Courtesy R. A. Page.)

occasions where the segment broke in two events separated by anything from 32 hours to three years². Generally, an earthquake along the eastern part of the trough triggered one to the west. Despite wide variations in the intervals between events, this history had been considered an example of characteristic (regular) earthquake behaviour. But important features of these events differ. For example, the 1605 event produced extensive tsunami damage, but no reports of damage by shaking as occurred at many of the other events (Kanamori).

Along the Nicaraguan trench, a very recent underthrusting earthquake (2 September 1992) produced an anomalously large and damaging tsunami for its surface-wave magnitude, but was barely

recent earthquakes have not occurred preferentially in the gaps, but have tended to cluster in both space and time (D. D. Jackson, University of California, Los Angeles). This behaviour may be partly due to clustering and partly due to aseismic slip along some plate boundaries, rendering them incapable of a large earthquake.

An example of the difficulty with the seismic-gap hypothesis can be seen in the Shumagin Gap, a section of the Aleutian subduction zone. Most of the Aleutian arc east and west of the Shumagin Island section has ruptured in great earthquakes this century (T. M. Boyd, Colorado School of Mines). From this history, the Shumagin section was designated a seismic gap and monitored intensely. However, a decade of geodetic surveil-

^{*}Wadati Conference on Great Subduction Zone Earthquakes, Fairbanks, Alaska, 16–19 September 1992.

ance shows that little strain is accumulating, compared with that to be expected were the two converging plates strongly locked together. The shortening of trilateration line lengths on the overriding plate in the direction of convergence is much less than expected for a mechanically locked plate contact, and subsidence is observed where a simple dislocation model predicts uplift (J. C. Savage, US Geological Survey, Menlo Park, California). Other measurements of vertical deformation also suggest that slip is mostly aseismic (J. Beavan, Lamont-Doherty Geological Observatory). Finite-element modelling with a visco-elastic rheology indicates that only about 15 per cent of the slip takes place seismically (R. Dmowska, Harvard University). This lack of accumulating strain casts doubt on the seismic potential of the Shumagin segment.

Perhaps the best hope for improving our understanding of earthquakes comes from integrated studies in particular regions. One promising area discussed at the conference is the Prince William Sound region, near the epicentre of the 1964 Alaskan earthquake, at magnitude 9.2 the second-largest recorded earthquake. The combined results of detailed seismicity studies, geodetic measurements and regional seismic sounding indicate that at least part of the megathrust (the major interplate boundary), which slipped tens of metres in 1964, is now seismically silent and presumably locked, accumulating strain. The distribution of slip in the 1964 event determined from bodywave seismograms (D. Christensen, University of Alaska, Fairbanks) agrees with that determined by inversion of geodetic data (S. R. Holdahl, National Geodetic Survey), and identifies two large regions of high slip beneath Prince William Sound and Kodiak Island.

Crustal structure in the Prince William Sound region has been imaged with seismic reflection and wide-angle reflection/refraction profiles from the Trans-Alaska Crustal Transect⁴ (TACT). A prominent, gently-dipping velocity jump, inferred from strong reflected and refracted arrivals, has been interpreted as the megathrust, in view of its depth, regional extent and relation to the location and stress orientation of seismicity in the downgoing plate. Carefully located seismicity, monitored with a regional network, occurs mainly in the upper 10 km of the subducting plate, whereas the megathrust is nearly aseismic (R. A. Page, US Geological Survey, Menlo Park). The observation that coseismic vertical deformations have been partially recovered in the post-earthquake period suggests that the interplate contact is locked⁵. The absence of aftershocks larger than magnitude 5.0

beneath Prince William Sound after the 1964 earthquake suggests that the interface locked quickly. The seismicity in the subducting plate reflects down-dip tensional deviatoric stress, consistent with a locked fault zone, and is spatially continuous with deep subduction-zone seismicity (Page). The seismicity, TACT results and regional tectonics indicate that the megathrust lies between the overriding plate and the subducted Yakutat terrane, which appears to be sutured to the top of the Pacific plate⁴ (see figure).

The seismic quiet on the Alaskan megathrust beneath Prince William Sound strengthens the case that the Cascadia subduction zone, 1,600 km further south, is also threatened by a great earthquake — a matter of concern for the 10 million people living in the region. Coastal geology strongly suggests that great Cascadia earthquakes have occurred in the past few thousand years (B. F. Atwater, US Geological Survey, Seattle). Moreover, strain measured geodetically suggests that a mostly offshore part of the plate boundary is now locked⁶. These findings have seemed inconsistent with the seismic record, which contains virtually no subduction interface earthquakes of any size in the twentieth century except for the Petrolia sequence in northern California in April this year. But the seismic quiet beneath Prince William Sound since 1964 shows that this absence of interplate earthquakes need not contradict the hypothesis of past great earthquakes on the Cascadia subduction zone. The geological evidence for these includes studies of coastal marshes submerged by regional coseismic downdrop and suggests significant variability in earthquake size and recurrence in Cascadia too (Atwater).

The variability that currently hinders earthquake forecasting may spur development of real-time warning systems for earthquakes and tsunamis⁷. Such systems might allow critical facilities, such as power plants, to be shut down moments before being struck by seismic or oceanic waves, and could help in rapid damage assessment afterwards □

Heidi Houston is at the Institute of Tectonics, Department of Earth Sciences, University of California, Santa Cruz, California 95064, USA.

1. Kanamori, H. & McNally, K. C. *Bull. seism. Soc. Am.* **72**, 1241–1254 (1982).
2. Ishibashi, K. in *Earthquake Prediction* (eds Simpson, D. W. & Richards, P. G.) 297–332 (1981).
3. Stein, S. *Nature* **356**, 387–388 (1992).
4. Brocher, T. M., et al. in *Continental Lithosphere: Deep Seismic Reflections* 241–246 (Am. Geophys. Un., 1991).
5. Savage, J. C. & Plafker, G. J. *Geophys. Res.* **96**, 4325–4335 (1991).
6. Savage, J. C. & Lisowski, M. *Science* **252**, 101 (1991).
7. Heaton, T. H. *Science* **228**, 987–990 (1985).

Warm feelings

Most people have an ambivalent attitude towards heat energy, especially when expressed in calories. They are great consumers of calories in fuel, but hate and fear them in food. Many householders grimly endure a thin, joyless starvation diet prepared in a tropically overheated kitchen.

Daedalus now points out the obvious escape from the contradiction. The householder should keep warm by generating more metabolic heat. This would simultaneously reduce heating bills and burn off the accumulating body fat. Sadly, the human body normally increases its metabolic rate only in response to thoroughly unpleasant stimuli, such as pain, cold or exercise. But Daedalus notes a minor benefit of alcohol: it dilates the blood vessels of the skin. It gives the drinker a warm glow, but by the same effect increases the rate at which heat is lost from his body. His internal reserves of energy can be depleted very rapidly. The Saint Bernard with the brandy cask is actually quite a hazard to the snowbound traveller.

Alcohol, of course, has drawbacks as well as benefits. It has a calorific content of its own, for example. So DREADCO biochemists are seeking a substance that brings the blood to the skin without supplying calories or making you drunk. They remark that alcohol is a slow-acting anaesthetic. Some other anaesthetics can also increase the metabolic rate. Certain high explosives, like dinitrophenol, also do this: and nitroglycerine is used as a drug to relax the arteries of coronary victims. The DREADCO team is accordingly combing the ranks of anaesthetics, high explosives, and even high-explosive anaesthetics, in search of the ideal heat-loss agent. Even if they don't find it, they may pick up enough clues to enable them to synthesize it. It will bring blood to the skin with a minimum of side-effects, giving the user a sense of warmth and comfort even in the coldest house, while burning up food calories at an unprecedented rate.

DREADCO's 'Body Heat'® should find a wide and grateful market. Its happy users will permanently radiate a warm, enticing, healthy glow. They will be able to tuck into a steady diet of fatty, sugary pies and pastries, while maintaining an enviably slim and elegant figure. They will be able to flaunt that figure in the thinnest and flimsiest of fashionable clothing without have to heat their environment to boilerhouse levels to stay comfortable. The only snag is that their raised metabolic rate may 'burn them out' sooner. But with the ranks of geriatrics already swelling ominously, even this may be an advantage.

David Jones

Thyroid cancer incidence

SIR — As mentioned in two recent contributions to Scientific Correspondence^{1,2}, the increase in the number of reported thyroid cancer cases in Belarus, particularly in Gomel^{3,4}, could have arisen from the effect of better reporting, heightened awareness and screening. To help interpret the Belarus results, we estimated the magnitude of the effect of screening in a cohort of children treated at the Michael Reese Hospital by irradiation for benign head and neck diseases.

In 1974, newspapers in the United States were drawing attention to the excess of thyroid cancer among people

two factors: (1) the intense screening efforts between 1974–79 detected early tumours and, therefore, fewer remained to be diagnosed later; and (2) screening was not as intense during subsequent years.

To reduce the effect of increased incidence from radiation exposure and thus isolate the effect of screening, a similar analysis was performed for the lowest-dose group category (less than 50 centigray). In this group, the difference in the incidence of thyroid tumours before and after 1974 was even greater (see table).

Our results were based on examination of adults; Mettler *et al.* have

INCIDENCE OF THYROID CANCER AND NODULES

Year of screening (All patients)	Cancer		All nodules*	
	No. cases	Rate 10 ⁻⁴ PY†	No. cases	Rate 10 ⁻⁴ PY
Before 1974	109	25.8	260	42.3
1974–79	169	191.9	625	726.2
During and after 1980	31	69.8	158	392.2
(Patients receiving <50 centigray)				
Before 1974	18	14.0	40	19.7
1974–79	47	169.1	138	418.9
During and after 1980	9	68.1	40	235.1

Patients were treated with external radiation for benign head and neck conditions at Michael Reese Hospital.

* Includes patients with malignant and benign tumours, as well as patients with thyroid nodules not surgically removed.

† Gender and age adjusted to a person aged 30 years. PY, person years.

treated with head and neck radiation. At the same time, Michael Reese Hospital began a recall and screening programme for more than 5,000 people who had received radiation treatment⁵. For 2,634 (68.5%) patients who were 15 years of age or younger at the time of treatment (mean age 4.3 years) and for whom thyroid doses were computed, information about the occurrence of thyroid nodules was obtained from a screening examination at the hospital or from a questionnaire (A.B.S. *et al.*, manuscript submitted). Surgically removed thyroid tumours reported in the questionnaires were histologically confirmed.

To evaluate the effect of screening, we compared incidence rates before the screening programme; during the most active period of screening (1974–79); and when the level of screening had subsided somewhat (after 1980). The 6-year period between 1974 and 1979 is most comparable, in terms of screening effects, to the period after the Chernobyl accident (1987–92). As seen in the table, the adjusted incidence rates were much greater during the 1974–79 period than before 1974 (7- and 17-fold for malignant and all thyroid nodules, respectively). These increases are of the order of those seen in Belarus. The decline in incidence rates seen after 1979 probably reflects

reported⁶ that 0.7% of a small sample of 5- and 10-year-olds living in contaminated and control settlements around Chernobyl had palpable thyroid nodules, suggesting that even among young children, thyroid nodules occur frequently and can be detected by screening. Although our data should not be interpreted as showing that there is no effect of radiation from Chernobyl, they demonstrate that carefully controlled epidemiological studies are needed to understand the true impact of the accident.

Elaine Ron

Jay Lubin

*Epidemiology and Biostatistics Program,
National Cancer Institute,
National Institutes of Health, Bethesda,
Maryland 20892, USA*

Arthur B. Schneider

*Division of Endocrinology and
Metabolism,*

*Humana and Michael Reese hospitals,
University of Illinois,
Chicago, Illinois 60616, USA*

1. Beral, V. & Reeves, G. *Nature* **359**, 680–681 (1992).
2. Shigematsu, I. & Thiessen, J. W. *Nature* **359**, 681 (1992).
3. Kazakov, V. S. *et al.* *Nature* **359**, 21 (1992).
4. Baverstock, K. *et al.* *Nature* **359**, 21–22 (1992).
5. Schneider, A. B. *et al.* *Medicine* **64**, 1–15 (1985).
6. Mettler, F. A. *et al.* *J. Am. med. Ass.* **268**, 616–619 (1992).

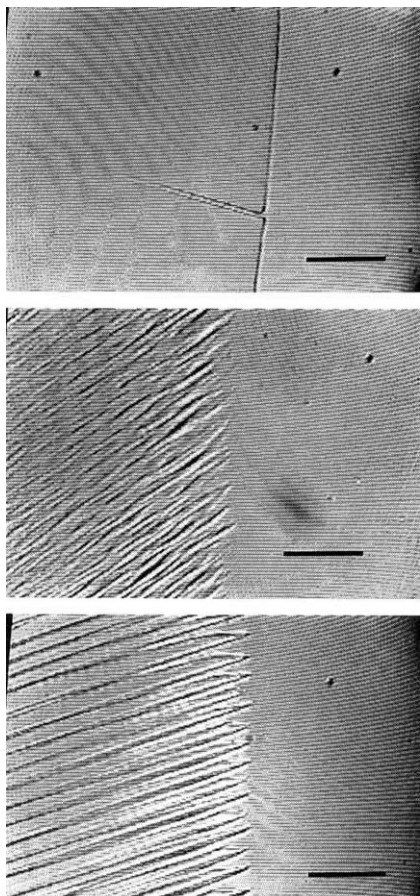
Ice-crystal growth and lectins

SIR — The C-type lectin found in rattlesnake venom can modify the growth habit of ice crystals in a manner similar to that previously thought unique to fish antifreeze proteins (AFP)^{1–3}. Four main types of AFP have been identified in the blood sera of various teleost fishes which protect fish from freezing by depressing the freezing temperature of their body fluids non-colligatively. These proteins act by preferentially binding to the non-basal planes of ice crystals, thereby modifying their growth habit to produce unique spicular, faceted structures rather than the usual smooth, non-faceted ones^{1–3}.

Recently, we reported that one of the AFP types (II) found in three distantly related fish species (sea raven, smelt and herring) has a structure similar to the carbohydrate-recognition domain (CRD) of calcium-dependent (C-type) lectins⁴. This prompted us to determine whether these lectins possess AFP-like properties. To this end, we used the lectin isolated from the venom of the rattlesnake (*Crotalus atrox*). This lectin is a dimer with a M_r of 28,000; the native form appears aggregated into a larger multimeric structure^{5,6}. Each monomer consists of a CRD. The lectin was purified from snake venom by affinity chromatography essentially following previously described methods^{5,6}.

Initial studies with this protein using a Clifton nanolitre osmometer did not reveal any non-colligative freezing point depression at the resolution of the device. However, using an apparatus that allows microscope examination of the freezing interface during unidirectional freezing under controlled temperature conditions (directional solidification stage⁷) we find that the lectin does possess properties that so far have been unique to antifreeze proteins^{1–3}. Selected results of unidirectional freezing, with freezing velocities of 2.5 ± 0.7 mm min⁻¹ and temperature gradients of 15 ± 1 °C mm⁻¹, are shown in the figure.

The upper panel illustrates, in two ice crystals separated by a grain boundary, the planar ice-crystal morphology typical of pure water and aqueous solutions containing low concentrations of solutes. This particular result was obtained with a 0.71 mM solution of NaCl in distilled water. A similar morphology was observed with distilled water; 2.5 mM and 5 mM CaCl₂; and 0.71 mM, 2.5 mM and 5 mM lactose. The smooth non-faceted morphology of the ice crystal suggests that the freezing interface is rough on an atomic or molecular scale, and that the growth rates are indepen-



Freezing is from left to right; isotherms are vertical; scale bar, 50 μm . See text for details.

dent of crystallographic orientation. The centre panel shows the spicular ice structure that forms in a 0.71 mM solution (a typical physiological concentration in fish) of sea raven antifreeze proteins in distilled water. This spicular structure is thought to result from the unique ability of antifreeze proteins to bind to the nonbasal planes of ice crystals and inhibit growth on these planes. This property has not been observed in other protein types³.

The lower panel obtained for a solution of 0.71 mM rattlesnake venom lectin in distilled water clearly shows faceted ice crystals. It is obvious that this ice-crystal morphology is not caused by the phenomenon of freezing interface instability, which is associated with rejection of solutes by ice, colligative depression of freezing temperature and constitutional supercooling, because the inorganic salt and the sugar controls

produced planar freezing interfaces at much higher concentrations. Furthermore, the cellular and dendritic structures caused by interface instability are smooth and non-faceted, indicating a rough interface on an atomic or molecular scale⁷. The ice crystals that form in the presence of the lectins are faceted, indicating that the ice-crystal facet seen is smooth on the atomic scale and is the slowest growing ice-crystal plane.

The results suggest that the lectins inhibit ice growth along this plane, a property thought to be unique to antifreeze proteins. The mechanism resulting in this effect is probably similar to that of the antifreeze proteins. The fact that only one crystallographic orientation is affected by the lectins, whereas the antifreeze proteins seem to affect all nonbasal planes, may explain why we did not observe any depression in freezing temperatures with lectins. However, if the ability to inhibit ice growth along particular crystallographic planes is com-

mon among all C-type lectin CRDs, then the CRDs are good prototype structures from which an AFP could have evolved. This may explain the evolution of type II AFP in three divergent fish groups. Furthermore, as the snake venom lectin, a protein that is not an antifreeze, appears to have the ability to modify the structure of ice crystals by interacting with specific ice-crystal planes, this property may be more common to proteins than had previously been thought.

B. Rubinsky

R. Cogger

Department of Mechanical Engineering,
University of California, Berkeley,
California 94720, USA

K. V. Ewart

G. L. Fletcher

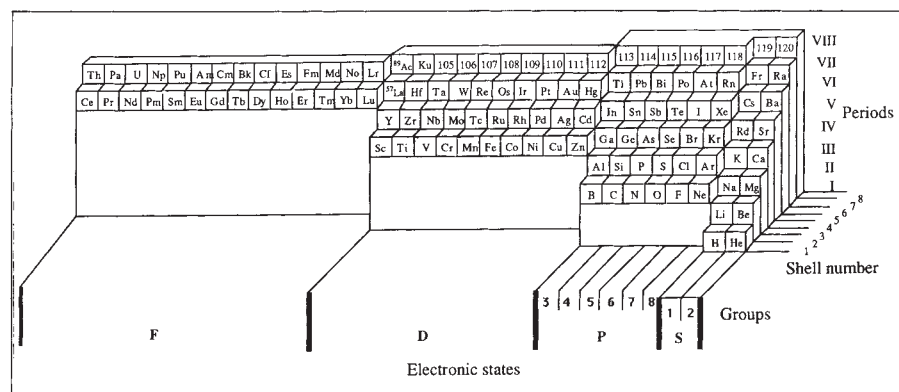
Ocean Sciences Centre and
Department of Biology,
Memorial University of Newfoundland,
St John's,
Newfoundland A1C 5S7,
Canada

A three-dimensional periodic table

SIR — More than 50 years ago, Linus Pauling suggested his classic idea of introducing another index, namely electronegativity of atoms, to the periodic table¹. Recently, Allen² introduced a third index into the periodic table, configuration energy (CE), and demonstrated its usefulness. As Maddox³ puts it, the question of whether "...CE resolves enough of the inconsistencies in the periodic table to win the hearts and minds of men with an interest in these matters" still remains to be

present, the second period consists of only Li and Be, the third has eight elements from B to Mg, the fourth again contains 8 elements from Al to Ca, and so on.

The three-dimensional picture, in which the third dimension is periodic number ($n+1$), is shown in the figure. All the elements on the same level have equal $n+1$. Such subdivision seems absolutely natural. The only obstacle is tradition. Since Mendeleev, every period begins with s -electrons of the previous



answered.

We propose a genuinely three-dimensional periodic table, based on quantum mechanics and Hund's rules (the latter proposed as early as 1926; ref. 4), namely, that the term of largest S and among these the term of largest L is lowest in energy. In this case the number of chemical elements in periods is the following: 2, 2, 8, 8, 18, 18, 32, 32, ...

Therefore, for instance, in period number 1 only hydrogen and helium are

level, so that it includes electrons from two levels, that is, s -electrons from the level $n-1$, and all the rest of the electrons from the level n . This mixing, from the point of view of quantum mechanics, is confusing.

Periods correspond to passing from one chemical element to another one at the same height from left to right. Group numbers are traditional ones. Two elements, La and Ac, are placed according to their traditional places, although they

1. Raymond, J. A. & DeVries, A. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2589–2593 (1977).
2. Knight, C. A., DeVries, A. L. & Oolman, L. D. *Nature* **308**, 295–296 (1984).
3. Harrison, K. et al. *Nature* **328**, 241–243 (1987).
4. Ewart, K. V., Rubinsky, B. & Fletcher, G. L. *Biochem. biophys. Res. Commun.* **185**, 335–340 (1992).
5. Hirabayashi, J., Kusunoki, T. & Kasai, K.-I. *J. biol. Chem.* **266**, 2320–2326 (1991).
6. Gartner, T. K. & Ogilvie, M. L. *Biochem. J.* **224**, 301–307 (1984).
7. Rubinsky, B. & Ikeda, M. *Cryobiology* **22**, 55–68 (1985).

could also be placed in the left-most position.

When projecting in the *x*-direction (number of shell), the three-dimensional table degenerates into two dimensions. When projecting down along the *z*-axis (axis of periods), we get a two-dimensional realization of Hund's rule.

In conclusion, we believe that our three-dimensional representation is a useful tool for visualizing properties of chemical elements and is in complete accord with quantum mechanics.

Y. Magarshak

J. Malinsky

Department of Mathematics,
Mount Sinai School of Medicine,
New York,
New York 10029, USA

1. Pauling, L. *The Nature of the Chemical Bond* (Cornell Univ. Press, 1939).
2. Allen, L. C. *J. Am. chem. Soc.* **114**, 1 (1992).
3. Maddox, J. *Nature* **356**, 13 (1992).
4. Hund, F. *Linienenspektren und Periodisches System der Elemente* (Berlin, 1927).

Eradication of rabies in Europe

SIR — To complete the story described by Brochier *et al.*¹, we would like to explain what had already been done in Europe since 1978, when oral vaccination of foxes against rabies began, as well as to mention a recent development concerning improvement of vaccinal strains.

Investigations of oral immunization of foxes began in Switzerland about 20 years ago, after US and Canadian workers had demonstrated that orally administered attenuated rabies strain SAD could immunize foxes². By 1978, a system was ready for field use, consisting of 1.8 ml of SAD (Bern) at a titre of 10^7 TCID₅₀ per ml (where TCID is tissue culture infectious dose), included in small crushable plastic containers (blisters) fixed under the scalp of chicken heads. The baits were manually distributed.

After a first successful trial on a 335-km² area, vaccination zones were gradually extended over the Swiss midlands and finally the Jura mountains bordering France. One after the other, the infected regions were freed from rabies by immunizing their fox populations through the distribution of 12–15 vaccine baits per km². The epidemiological effect of vaccination was most dramatically demonstrated in eastern Switzerland, where high rabies prevalence lasting for 18 years came to a sudden end after only three vaccination campaigns covering the whole region³.

Five years after the first trial in Switzerland, Germany began to use the same method with SAD B19, a deriva-

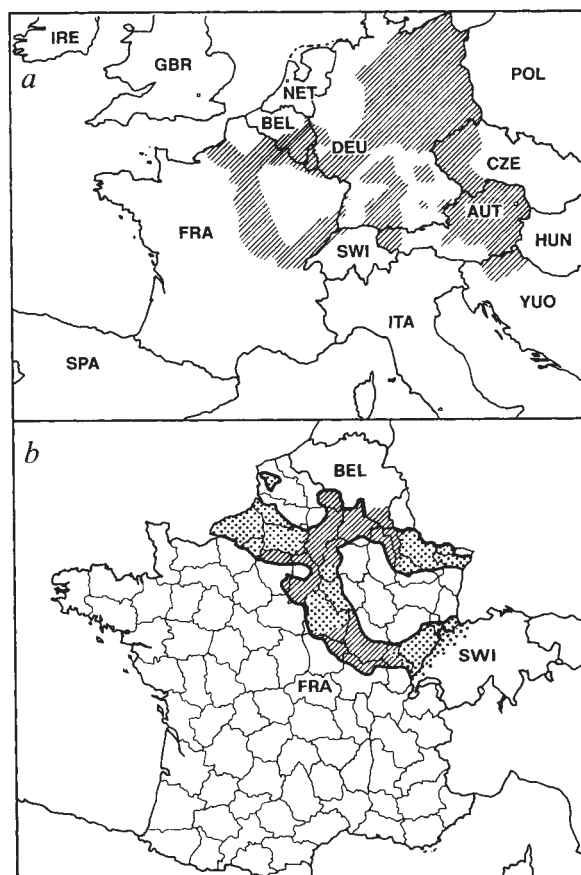
tive of the Swiss vaccine strain. In 1985 the Germans developed an 'artificial bait' based on tallow and fish meal that could be mechanically produced⁴, which considerably simplified the preparation, storage and distribution of the vaccine.

After 1983, field trials were begun in several other European countries and in Canada. In Europe alone, a total of around 300,000 km² was treated in the spring and autumn of 1991 (*a* in the figure) representing about 9 million baits dropped by planes and helicopters or distributed by hand, leading to a consistent reduction in the number of cases of rabies.

Although poorly virulent by intramuscular or oral inoculation routes, fixed strains used for vaccination of foxes exhibited residual pathogenicity, at least for rodents⁵, and this was a matter of concern for the authorities.

Previous work on the molecular basis of rabies virulence has shown that avirulent mutants could be isolated from virulent strains, using appropriate monoclonal antibodies⁶. We isolated several of those antibodies that neutralize SAD (Bern) or derivatives but do not neutralize mutants with a substitution in position 333 of the viral glycoprotein, a mutation which abolishes viral virulence for adult mice^{7–9}. One of those monoclonals was used to select SAG1, a mutant of SAD (Bern) which had a serine in position 333 instead of an arginine and, as expected on the basis of previous results, was avirulent for adult mice, whatever the dose and the route of inoculation. When administered orally or by the mucosal route it also failed to induce any pathology in foxes, in non-target carnivora (badgers, ferrets, polecats), in crows and birds of prey, and in several species of wild rodents¹⁰.

SAG1 strain was found to be as immunogenic as the parental strain: 100% survival was achieved for doses equal to or higher than 10^6 plaque-forming units given orally to foxes. Lower doses still induced some protection. This mutant was found to be fairly stable and did not revert to virulence after three successive intracerebral inoculations of suckling mice, unlike a similar avirulent mutant (Arg→Ser) isolated from another fixed



Oral vaccination of foxes against rabies. *a*, In Europe. Hatched area; vaccination campaigns in 1991 (subject to a few possible modifications not yet registered by WHO). (Adapted from *The work of WHO 1990–1991: Biennial Report of the Director General* by permission of the World Health Organisation, which retains the copyright.) *b*, In France (FRA), Belgium (BEL) and Switzerland (SWI). Dotted area, vaccination with the Virbac vaccine; hatched area, vaccination with the Rhône-Mérieux vaccine.

strain of rabies, CVS (ref. 8). The reason for this difference is not clear, but it could be due either to a lower mutation rate of the SAD (Bern) polymerase or to a lower selective advantage of the revertants in this genetic background. More recently, a further step toward complete safety of the vaccine was accomplished by the selection of a double avirulent mutant, SAG2, which has two mutations in the codon of Arg 333. This mutant is under trial at present.

In France, the SAD B19 has gradually

1. Brochier, B. *et al. Nature* **354**, 520–522 (1991).
2. Debbie, J. G., Ableseth, M. K. & Baer, G. M. *Am. J. Epidemiol.* **96**, 231–235 (1972).
3. Wandeler, A. I., Capt, S., Kappeler, A. & Hauser, R. *Rev. Infectious Diseases* **10**, S649–S653 (1988).
4. Schneider, L. G. *et al. Rev. Infectious Diseases* **10**, S654–S659 (1988).
5. Leblois, H. & Flamand, A. in *Vaccination to Control Rabies in Foxes* (eds Pastoret, P. P. *et al.*) 101–104, Office for Official Publ. Eur. Commun. Eur 11439 (1988).
6. Coulon, P., Rollin, P., Aubert, M. & Flamand, A. *J. gen. Virol.* **61**, 97–100 (1982).
7. Seif, I., Coulon, P., Rollin, P. E. & Flamand, A. *J. Virol.* **53**, 926–934 (1985).
8. Tuffreau, C. *et al. Virology* **172**, 206–212 (1989).
9. Dietzschold, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 70–74 (1983).
10. Artois, M. *et al. Vaccine* **10**, 524–528 (1992).

been replaced since 1990 by other commercially available artificial baits containing new generations of vaccines: the previously described SAG1 strain developed by the CNRS and Virbac, and the recombinant vaccinia virus containing the rabies glycoprotein gene (Rhône-Mérieux)¹. Vaccination is extended to the whole contaminated area (112,000 km²), half with each virus (*b* in the figure). In Switzerland, SAG1 vaccine has been used since spring 1991, whereas recombinant vaccinia has been chosen in Belgium. Other European countries still use classical fixed strains of rabies.

A. Flamand

P. Coulon

F. Lafay

Laboratoire de Génétique des Virus,
CNRS, 91198 Gif sur Yvette Cedex, France

A. Kappeler

Swiss Rabies Centre,
Institute of Veterinary Virology,
University of Bern,
3012 Bern, Switzerland

M. Artois

M. Aubert

Laboratoire d'Etudes sur la Rage,
CNEVA, BP9, 54220 Malzeville, France

J. Blancou

Office International des Epizooties,
12, rue de Prony, 75017 Paris, France

A. I. Wandeler

Agriculture Canada,
Animal Diseases Research Institute,
PO Box 11300, Station H,
Nepean, Ontario K2H 8P9, Canada

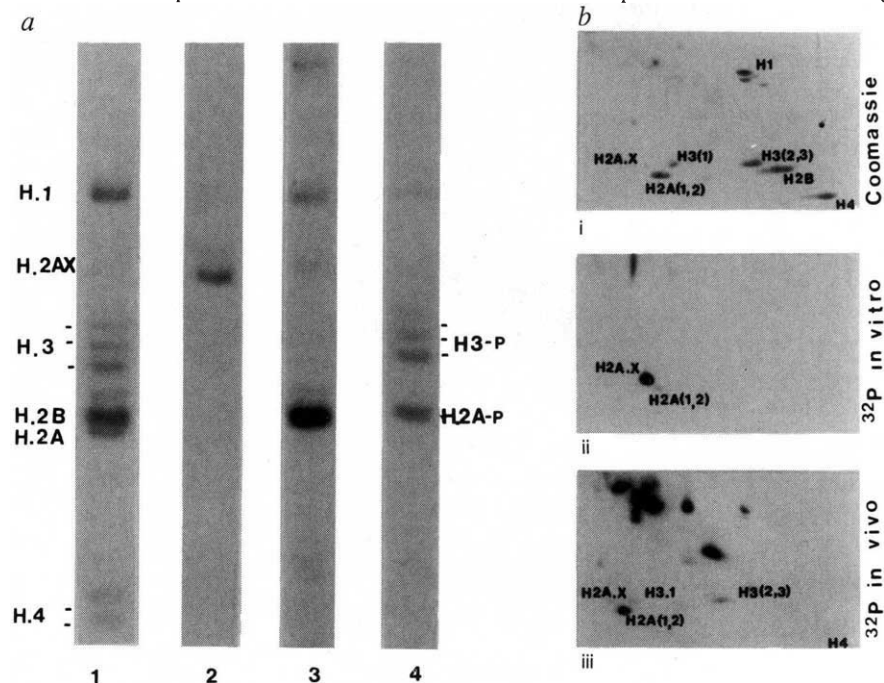
Which histone kinase?

SIR — Growth factors, phorbol esters, okadaic acid and protein-synthesis inhibitors stimulate rapid phosphorylation of the nucleosomal core histone H3 (ref. 1), the duration and strength of which correlates well with induction and super-induction of the oncogenes *c-fos* and *c-jun*. To identify possible kinase(s) involved in this response we have repeated the experiments of Whitlock *et al.*², who described the existence of a chromatin-associated, calcium-dependent histone H3 kinase whose activity is enhanced in butyrate-treated cells. Although the phosphorylation phenomenon they described is entirely reproducible, the phosphoprotein involved is not histone H3, but a minor, then undiscovered, variant of H2A called H2A.X.

We prepared nuclei from quiescent and stimulated C3H 10T1/2 cells, and performed *in vitro* kinase assays using conditions identical to those described in ref. 2. Irrespective of whether cells were stimulated or not, the main ³²P-labelled phosphoprotein migrates just above the

maximally acetylated forms of histone H3 on acid-urea gels (*a* in the figure and Fig. 2 of ref. 2). In agreement with Whitlock *et al.*², we find that this phosphoprotein comigrates exactly with histone H3 in SDS/PAGE gels and its phosphorylation is augmented in the presence of calcium and in nuclei derived from cells pretreated with sodium

produces marked phosphorylation of the minor H2A.X variant, but negligible phosphorylation of H2A.1 and H2A.2 despite the large stoichiometric excess of these latter species⁵. Although most of the work on H2A.X phosphorylation has been done using *in vitro* models, DNA-dependent phosphorylation of H2A.X has been implicated in modulating



a, Acid-urea gels comparing *in vitro* and *in vivo* ³²P-labelled histones. Lane 1, Coomassie staining; lanes 2–4, autoradiograms of *in vitro*-labelled histones (lane 2), histones labelled *in vivo* from unstimulated cells (lane 3) and cells stimulated with 50 ng ml⁻¹ epidermal growth factor, plus 10 μg ml⁻¹ anisomycin for 1 h (lane 4). *b*, TAU/SDS two-dimensional electrophoresis⁴ of the above. *i*, Coomassie blue staining; *ii*, ³²P-ATP *in vitro*-labelled histones; *iii*, ³²P-phosphate *in vivo*-labelled histones from stimulated cells.

METHODS. Cells were washed twice with ice-cold PBS and nuclei prepared either by lysis in 10 mM HEPES pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM dithiothreitol and 1% NP-40 (ref. 7), or exactly as described in ref. 2 with similar results. Nuclei were washed, ³²P-ATP-labelled and histones prepared as described². ³²P-labelling and cell stimulation *in vivo* is described in ref. 1.

butyrate^{2,3}. But this phosphoprotein did not comigrate on acid-urea gels (*a* in the figure) with phosphorylated histone H3 isolated from mitogen-stimulated cells labelled *in vivo* with ³²P-phosphate¹.

To resolve this problem we used a two-dimensional gel system, with Triton-acid-urea (TAU) gels⁴ in the first dimension and SDS gels in the second, capable of resolving variant forms of the core histones (*b* in the figure). In particular, the inclusion of 0.37% Triton X-100 in one-dimensional gels allows unambiguous resolution of the main variants of histones H3 and H2A. This shows that the main *in vitro* phosphorylated species observed here and in ref. 2 is not histone H3 but H2A.X (ref. 5). Its mis-identification arises from its precise comigration with histone H3 in SDS gels and its migration close to hyperacetylated H3 in acid-urea gels.

It is intriguing that the *in vitro* kinase activity described here and in ref. 2

nucleosome spacing in *Xenopus* oocytes⁶. We thank Ron Laskey (Wellcome/CRC Institute, Cambridge) for suggesting the possible identity of this phosphoprotein.

Eva Cano

Michael J. Barratt

Louis C. Mahadevan

Nuclear Signalling Laboratory,
Developmental Biology Research Centre,
King's College London,
26–29 Drury Lane,
London WC2B 5RL, UK

1. Mahadevan, L. C., Willis, A. C. & Barratt, M. J. *Cell* **65**, 775–783 (1991).
2. Whitlock, J. P. Jr, Augustine, R. & Schulman, H. *Nature* **287**, 74–76 (1980).
3. Whitlock, J. P. Jr, Galeazzi, D. & Schulman, H. *J. Biol. Chem.* **258**, 1299–1304 (1983).
4. Alfageme, C. R., Zweidler, A., Mahowald, A. & Cohen, L. H. *J. Biol. Chem.* **249**, 3729–3736 (1974).
5. West, M. H. P. & Bonner, W. M. *Biochemistry* **19**, 3228–3245 (1980).
6. Kleinschmidt, J. A. & Steinbeisser, H. *EMBO J.* **10**, 3043–3050 (1991).
7. Mahadevan, L. C., Willis, A. J., Hirst, E. A., Rathjen, P. D. & Heath, J. K. *Oncogene* **5**, 327–335 (1990).

Vagaries of thought

Christopher Longuet-Higgins

A History of the Mind: Evolution and the Birth of Consciousness. By Nicholas Humphrey. Chatto and Windus/Simon and Schuster: 1992. Pp. 238. £16.99, \$22.00.

The Strange, Familiar and Forgotten: An Anatomy of Consciousness. By Israel Rosenfield. Knopf/Picador: 1992. Pp. 157. \$20, £14.99.

THERE has been quite a kerfuffle about consciousness recently. Books have appeared on the subject by, among others, Daniel Dennett, philosophical spokesman for the artificial-intelligence community, and John Searle, its most formidable critic*. Not long ago, Roger Penrose invoked the fact of consciousness as evidence against the idea of artificial intelligence†. In June this year, the Wellcome Foundation awarded consciousness a certificate of scientific respectability in the form of a two-day international meeting in London. At the beginning of August, Dennett appeared on BBC television to announce that after centuries of scientific and philosophical fumbling we at last had a theory of "the phenomenon of consciousness", namely his own. In the September issue of *Scientific American*, Francis Crick and Christof Koch claim that "the problem of consciousness" is on the verge of solution, but that Dennett's theory is not quite right. Now Nicholas Humphrey, biologist and communicator, speculates on the nature and evolution of consciousness, and Israel Rosenfield, a neurologist of the school of Oliver Sacks, ventures a slim volume on its anatomy. As Humphrey remarks disarmingly in his preface:

There have been not a few — perhaps too many — books on mind, consciousness and evolution published in the last few years (two of them by me). And, as shelves sag and appetites fade, I should explain what is different about this one.

One unusual feature of *A History of the Mind*, as Humphrey warns, is the little that it has to say about computers, artificial intelligence or quantum mechanics. But even more distinctive is the author's evident respect for his subject matter, informed by a fine literary sensibility. The book is exceptionally well written and its arguments attractively presented, even if they sometimes fail to carry conviction.

Humphrey summarizes his central

thesis in the aphorism "I feel, therefore I am." Turning the tables on Descartes — and on his own earlier views — he identifies the birth of consciousness with the advent of "raw sensation" as opposed to sensory perception or rational thought. This is a good point. We deprive people and animals of consciousness — by the use of anaesthetics — in order to save them from *feeling* anything, not from seeing or thinking. Humphrey makes a persuasive distinction between perception and sensation; but he then attempts to establish that perception can take place without sensation, dredging up 'skin vision' and 'blindsight' — both a long way from common experience — as relevant evidence. My unease at this line of thought became acute at the top of page 162, where the author postulates that not one but two processes accompany "the activity of sensing". One is something called "sentition", occurring centrally; the other is a barrage of "sentiments" or actual physical events at the body's surface, triggered by

effluent signals. A critical reader might well switch off at this point; or if not here, on encountering the author's U-turn on page 173:

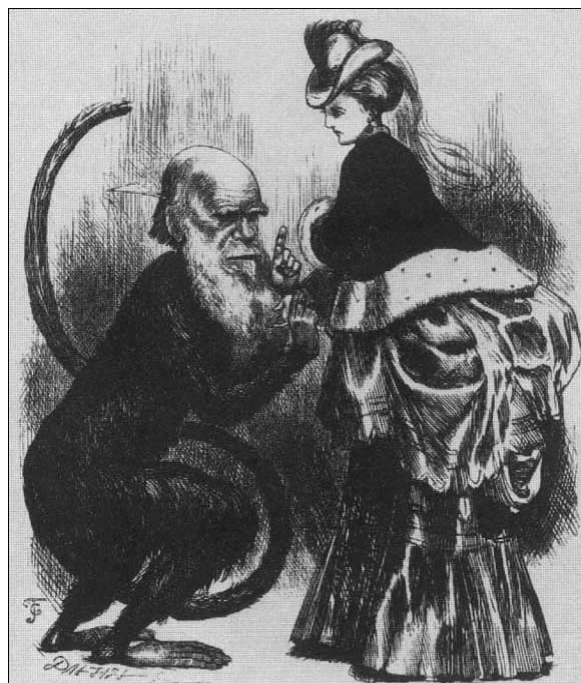
If the Mark-1 theory cannot cope, we need a Mark-2 theory that, while retaining the essential features of the previous version, is better adapted to the facts.

Mark 2 is, mercifully, a bit better, as it requires only that sensations, to qualify as such, be registered neurally in some location associated with the bodily location from where they originate (if I take Humphrey's meaning correctly). This location he identifies as the body's image in the cerebral cortex, thus attributing animal consciousness to re-entrant neuronal activity. Unfortunately the theory is less than specific about the sufficient conditions for consciousness to arise in creatures other than people and animals. What about plants, for example? Or robots? Conceding that it might be possible in principle for an engineer to construct a conscious machine, Humphrey dismisses this as impractical on the arguably irrelevant grounds that

there would be no way of re-creating the natural *historical traditions* that have given the activity occurring in natural brains the peculiar *modal quality* of consciousness.

(Could it have been an awareness of historical tradition that moved him to appear on the dust cover in the likeness of a mediaeval saint?)

Rosenfield's book *The Strange, Familiar and Forgotten* starts well, stressing



Femine flushes — the caption reads: "Really, Mr. Darwin, say what you like about man; but I wish you would leave my emotions alone!" In *The Expression of the Emotions*, Darwin wrote that "a pretty girl blushes when a young man gazes intently at her" because she immediately thinks about the "outer and visible parts" of her body, and this alters their "capillary circulation". The cartoon is taken from Adrian Desmond and James Moore's *Darwin*, now published in paperback by Penguin (price £9.99). "Unquestionably the finest biography ever written about Darwin", wrote Stephen Jay Gould in his review in *Nature* 355, 215 (1992).

* Daniel Dennett's *Consciousness Explained* was reviewed in *Nature* 356, 26 (1992); John Searle's *The Rediscovery of Mind* will be reviewed in *Nature's* Autumn Books supplement (26 November).

† See the review of Roger Penrose's *The Emperor's New Mind* in *Nature* 341, 393 (1989).

an attribute of consciousness that Humphrey neglects, namely its role in the exercise of memory. The ability to recall a particular event is wholly dependent on having been conscious at the time the event took place. The laying down of memories and their subsequent retrieval are processes of obvious biological usefulness, and insofar as consciousness is a precondition for them to occur, that should be excuse enough for its evolutionary appearance.

As a neurologist, Rosenfield has firsthand knowledge of amnesic disorders such as Korsakoff's syndrome, which combines a loss of memory for recent events with the retention of remote memories — to put the matter in ordinary language. But Rosenfield wants to go deeper:

To view such a neurological disorder as simply a matter of memory loss is to miss its much greater significance ... our memory of events, people and situations experienced in the remote past, and their relation to our present self, is quite different from our view of events a week, a month, or even a year ago ... Remote memory ... and recent memory are different modes of thought ...

But such declarations are not without their logical implications. If someone loses his or her long-term memory in a road accident, are we to infer that the accident just happens to coincide with a sea change in that person's "modes of thought"?

Like Sacks, who gives the book unqualified praise on the dust jacket, Rosenfield is at his best in describing the vagaries and dilemmas of real patients. But it is not without interest that from a rather different perspective he arrives at a position similar to Humphrey's in assigning the body image a central role in the maintenance of consciousness:

All memories are subjective — a subjectivity that the brain establishes by organizing all stimuli in terms of a dynamic, ever-changing 'body image', the brain's unique frame of reference. When brain damage deforms consciousness — that is, when it changes the brain's capacity for sustaining the body image and creating memories, one's very self is changed too.

The book is nicely printed, with generous margins, but if one opens it too widely it falls to pieces. □

Christopher Longuet-Higgins is in the Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, UK.

Addendum

The editors-in-chief of *AIDS* 1991, reviewed in *Nature* 359, 491 (1992), are M. W. Adler, J. W. H. Gold, J. E. Groopman, J. A. Levy and J. N. Weber. F. Wong-Staal edited the section on virology only, together with J. Goudsmit.

Original thinking

Andrew Whiten

Animal Minds. By Donald R. Griffin. University of Chicago Press: 1992. Pp. 320. \$24.95, £19.95.

A GREEN-backed heron breaks a twig into small pieces and uses it as bait to aid fishing: casting it onto the water, she watches it intently until an unfortunate minnow investigates. Her behaviour is immensely flexible. She may drop the bait from an overhanging perch or from the water's edge, retrieving it if it floats away. Describing this example along with a wonderful array of equally fascinating studies of animals' behavioural versatility, Griffin is drawn to two broad conclusions: that the minds of such animals must incorporate considerable cognitive complexity, and that often this will involve conscious, if simple, thinking.

As a review of findings supporting the first of these two conclusions, Griffin's book has no equal in the sheer range of species, studies and behavioural phenomena that he draws together. The originality of his contribution lies very

much in the sheer scope of what he is prepared to consider. For example, unlike authors of most books on animal cognition, he devotes long tracts to the sophistication of invertebrate behaviour, and gives more space to the diversity revealed by naturalistic, ethological studies than to traditional comparative psychology; yet he touches on virtually every important issue in the contemporary debate about animal minds. For the student of animal cognition, this is an excellent, up-to-date source book.

But it is in the second main proposition — that much of animal cognition implies conscious thought — that Griffin really sticks his neck out. It is not the first time he has done so. Famed originally for his discovery (with R. Galambos) of echolocation in bats, Griffin clearly has an extraordinary facility for insight into the perceptual worlds of other species, and in his earlier books *The Question of Animal Awareness* (Rockefeller University Press, 1976) and *Animal Thinking* (Harvard University Press, 1984) he developed a preoccupation with the question of what it is like to be a bat (or bonobo, or bee ...). This goal gained few adherents and elicited a lot of critical reaction (much of which Griffin boldly quotes and dismisses with relish). Nevertheless, it is likely that many who have been prepared to



The greater adjutant stork *Leptoptilos dubius*, so called for its supposedly military gait, is the most massive and ugliest of the Asian storks. It is also one of the most endangered, although in India the species has rarely been killed for food — as a carrion-eater, it is considered by Hindus to be unclean and therefore inedible. This drawing by D. Quinn is taken from *Storks, Ibises and Spoonbills of the World*, which claims to be the definitive monograph on these birds; the book is written by J. A. Hancock, J. A. Kushlan and M. P. Kahl and illustrated also by A. Harris. Academic, £65, \$139.

crew a flagship of cognitive ethology into uncharted waters in recent years have done so with more imagination and assurance as a result of seeing a scientist of such repute straining forward on the prow, confidently scanning the mists ahead for opportunities to explore the far frontier of animal consciousness.

Yet many, probably most, of these ethologists still take an agnostic position on the possibility of a science of animals' phenomenal awareness and they share this attitude with commentators and workers in related disciplines, notably philosophy and experimental psychology: crudely put, animal cognition, but not animal consciousness, is currently fair game for science. This is despite the fact that, in the same way that progress in the general field of cognitive science has prepared the ground for studies of animal cognition, human consciousness has recently become the subject of a burgeoning field of 'respectable' scientific enquiry, from neurobiology to computational psychology. Cognitivists are not, as Griffin so often implies, dismissing the likelihood that conscious processes are at work in animals. They are simply at a loss in scientifically distinguishing the conscious and the nonconscious in nonhumans.

The new book offers little to change this state of affairs. Griffin's typical approach is to describe in alluring detail some elaborate behavioural system such as beaver engineering, which clearly involves the integration of a diversity of information, complex decision-making and a flexible repertoire of adaptive behaviour — in short, a cognitive apparatus that begs investigation. But then he takes a mental leap, saying, in effect, "isn't it likely that this cognitive complexity entails conscious thoughts and feelings?" Maybe, but Griffin offers no operational criteria by which to establish this, nor does he show the reader how establishing that a particular cognitive operation is conscious would have any further implications for understanding animal minds. The repeated and, to my mind, repeatedly ineffectual wavings of the consciousness flag become rather tedious, but with practice these typically end-of-paragraph pleadings can be anticipated and skirted.

A more serious criticism is that with respect to the two main themes — cognition and consciousness — there is an unsatisfying lack of unifying models to tie together the profusion of specific studies and findings, descriptions of which are the book's strength. So, on cognition, we are never presented with anything approaching an integrative model of the cognitive system as a whole for any of the species or behaviour systems discussed. And on consciousness, it is equally disappointing to find a

biologist so concerned with the existence of this phenomenon offering no theory of what distinguishes consciousness as a biological process, nor suggestions as to what its function might be, despite the considerable growth of literature on these questions from perspectives as diverse as those of Francis Crick, John Crook, P. N. Johnson-Laird and Tim Shallice. □

Andrew Whiten is in the Department of Psychology, University of St Andrews, St Andrews, Fife KY16 9JU, Scotland, UK.

Algae in red

Ralph A. Lewin

Dunaliella: Physiology, Biochemistry and Biotechnology. Edited by M. Avron and A. Ben-Amotz. CRC Press: 1992. Pp. 240. £109, \$149.95.

FLY over San Francisco Bay on a sunny day and look down over the coastal lagoons, and you may see some that are green and others that are orange or red. Similar ponds can be found along the coasts of Chile, China, Israel and Australia, and in inland saline areas in, for example, Utah in the United States and central Asia. It is a fairly safe bet that you are looking at dense growths of a single-celled, naked biflagellate called *Dunaliella salina*. This must be the only alga that one can identify specifically from a height of several thousand kilometres. It is also the only algal species that can grow in salt solutions strong enough to pickle cucumbers and that can survive in saturated brine. The cells produce large amounts of red intracellular pigment globules when stressed by high insolation, high salinity and a deficiency of nitrogen or phosphorus. The redness of salt ponds along the Mediterranean coast of France was first attributed to a flagellate by Dunal in 1837, and later confirmed by a special commission sent by the Academy of Sciences in Paris. The alga was eventually named after Dunal by E. C. Teodoresco in 1905.

But why does it now merit a whole book? (Indeed, this is the second text on the subject, although the first* was published in Russian and is not readily accessible in the West.) In short, it is because there is a market for the orange pigment β -carotene, which *Dunaliella* cells can accumulate in large amounts (up to ten per cent of their dry weight) when stressed. It is commercially pro-

duced in Australia, China, Israel, the United States and perhaps also in the Ukraine. Ami Ben-Amotz and the late M. Avron, both long involved with *Dunaliella* biotechnology in Israel, have edited a volume that is relevant and timely. With a dozen authors, most from Israel but five from Europe, Japan and the United States, the book covers various aspects of the biology, cultivation and use of *Dunaliella*.

But first a little about what kinds of species are included under the generic name *Dunaliella*, because they are too dissimilar to permit generalizations. There are about two-dozen recognized species. A few, all rare, occur in non-saline waters. One, *D. (Spermatozopsis) acidophila*, from sulphurous volcanic springs, can be readily grown in 1 per cent sulphuric acid if suitably illuminated and supplied with sources of nitrogen and phosphorus (chapter 5). At pH values of 3 or above, it dies. A handful of species, all rather similar, are members of the marine phytoplankton; several cultured strains are being used as larval food in mariculture. And the rest, all obligate halophiles, include two or three that go red, and a lot more that remain green like typical chlorophytes. It is chiefly the red ones that are of economic importance.

At least two sources of confusion are not clarified in this book. Unless it is specified that the studies have been made with pure clones, they may include ambiguous results based on mixtures of true *D. salina* with evergreen 'weed' species such as *D. viridis*. And (for reasons not unconnected with patriotism and patent law) the editors have chosen (though not consistently) to refer to one high-carotenoid strain of *D. salina* by a *nomen nudum*, "*D. bardawil*" (after the site where it was originally collected). *Caveat lector!*

Chapter 1, on morphology and taxonomy, summarizes some of the contents of Masjuk's book. Chapter 2, on motility, regrettably makes no reference to some exhaustive, though old, studies published by Teodoresco in 1938 on bioconvection, which could be of practical importance for collecting the cells. Chapter 3 discusses photosynthesis. Chapter 4 deals with ion transport (of obvious importance, considering how well cells of some species can exclude hydrogen or sodium ions). Chapter 6 is on osmoregulation, which is effected by synthesis of intracellular glycerol when the salinity is high and by letting it out when the salinity falls. Chapter 7 reviews NMR studies by an enthusiastic exponent of this technique, whereby one can determine information on cell contents *in vivo*. Chapter 8 on molecular biology, a *sine qua non* today, foreshadows future technological developments. Chapter 9 is

* N. P. Masjuk's *Morphology, Taxonomy, Ecology and Geographic Distribution of the genus Dunaliella Teod. and Prospects for its Potential Utilisation* (Dumka, Kiev, 1973).

perhaps the most important, because it deals with carotene synthesis. And, finally, chapter 10 presents data on the possible value of the species in nutrition and therapy.

Carotene (from *Dunaliella*) is an apparently harmless yellow pigment, a precursor of vitamin A, an anti-oxidant and perhaps an anti-cancer agent (for which encouraging evidence seems to be accumulating). But here I should mention another caveat. There is some question as to whether 'natural' β -carotene (rich in the 9-*cis* form found in

Dunaliella) is nutritionally superior to the synthetic product (all *trans*, and sold at a considerably lower price on the world market), because the natural form tends to isomerize.

Despite the shortcomings I have mentioned, the book is a valuable contribution to the psychological and biotechnological literature. It is well printed, lavishly bound and expensive. □

Ralph A. Lewin is in the Scripps Institution of Oceanography, University of California, San Diego, California 92093-0202, USA.

Evolution of good taste

Tristram D. Wyatt and Martin C. Birch

Insect Chemical Ecology: An Evolutionary Approach. Edited by Bernard D. Roitberg and Murray B. Isman. *Chapman and Hall*: 1992. Pp. 359. £55, \$75 (hbk); £24.95, \$35 (pbk).

WHILE the songs of crickets and grasshoppers have provided biologists with some of their best tools for investigating central evolutionary questions, the largely chemically mediated world of most insects has hardly been touched. This is a pity. Few other animal groups can match the intensity of study given to insect sex pheromones, and as P. L. Phelan suggests, this offers exceptional opportunities to test important questions about, say, the evolution of reproductive isolation, speciation and the relative contribution of sexual selection. The sensory world of most insects is dominated by smell and taste and is involved in finding mates, oviposition sites, food plants and more. In each of these areas there are increasingly well-studied systems ready to be explored by the evolutionary biologist. Evidence for testing conflicting hypotheses awaits recognition.

Genetic engineers splicing genes for bacterial toxins into crop plants to provide resistance to insect herbivores may also be interested. For, as M. A. Caprio and B. E. Tabashnik note, "the remarkable diversity of [natural] plant defensive compounds is matched only by the ability of insects to overcome them". Will defences derived from *Bacillus thuringiensis* prove to be different?

Evolutionary biologists are one target audience for the book, but more important perhaps are chemical ecologists themselves. Traditionally, insect chemical ecology has focused on the chemicals involved and the immediate questions of the function of behaviours such as mate location by sex pheromones. Evolutionary questions have too often been ignored. Niko Tinbergen's plea to integrate the "four whys" (function, caus-

ation, phylogeny and development) has never been more needed, and is a very welcome theme throughout this book.

Our growing knowledge about insect chemical ecology is largely due to three developments: (1) greatly improved techniques for analysis, which have done for pheromones and other chemicals what tape and sonography did for studies of sound (although producing the 'playback' can still be a problem — insects and plants are often better and more precise chemists than we are); (2) a greater knowledge of the sensory physiology of insects; and (3) a more evolutionary approach.

This book is a timely reminder of how productive an evolutionary perspective can be, providing explanations of otherwise baffling problems. Among the topics covered are the evolution of plant-insect interactions, semiochemicals and insect sociality, pheromones and mating systems, and the sensory physiology of chemical perception.

In the last chapter, J. N. McNeil shows how pure and applied scientists, often mutually distrustful and dismissive, could gain from working together. For example, the behavioural ecologists' understanding of alternative mating strategies can help to explain the results of attempts by applied entomologists to use synthetic sex pheromones to disrupt the mating of moths in crops; ecological insights could make pest control more effective; and evolutionary biologists could exploit the knowledge gained from intensive study of pest species.

Chemical ecologists will find this book an invaluable stimulus, encouraging them to pose questions in a more evolutionary and ultimately a more productive way. □

Tristram D. Wyatt is in the Departments of Continuing Education and Zoology, and Martin C. Birch is in the Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

Stellar theory

R. J. Tayler

Structure and Evolution of Single and Binary Stars. By C. W. H. de Loore and C. Doom. *Kluwer*: 1992. Pp. 458. DFL 240, £84, \$142 (hbk); DFL 120, £42, \$69 (pbk).

OVER the past 20 years, astronomers have come to realize that considerable mass is lost from stars at many phases in their evolution and that many of the most dramatic stellar events involve the interaction of two stars in a close binary system. Most textbooks on stellar evolution are largely concerned with the properties of single, spherical, constant-mass stars and are therefore not equipped for a discussion of mass loss and binary evolution. But de Loore and Doom have been personally involved in these fields so it is natural that these aspects feature prominently in their book.

The authors' theoretical discussion of the equations of stellar structure and of the physics of stellar interiors is less rigorous and detailed than that of another recent text, *Stellar Structure and Evolution* by R. Kippenhahn and A. Weigert (Springer, 1990). The main emphasis of de Loore and Doom is on the results of calculations of the evolution of stars of a wide range of masses. When there is uncertainty about the validity of theoretical models, as in the phenomenon known as convective overshooting, they discuss a variety of calculations that indicate how important the uncertainty is. In addition, they tabulate the results of many calculations of the evolution of both single and binary stars. For binary stars, they have also displayed information about the observed properties of many systems. These features will make the book a useful source.

Unfortunately, the text has not been checked very carefully and there are many typographical errors. The ones I noticed should not trouble readers already familiar with the field but will cause some problems for students. More serious is that there seems to be an error in one of the fundamental equations of stellar structure. The authors' energy equation differs from that in other standard texts by including an explicit term for the rate of change of gravitational energy, whereas another term in the equation, the rate of work done by the pressure, includes the change of gravitational energy implicitly. As a result, the set of equations does not conserve energy in rapid evolutionary phases. □

R. J. Tayler is at the Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK.

Gene-for-gene coevolution between plants and parasites

John N. Thompson & Jeremy J. Burdon

The concept of gene-for-gene coevolution is a major model for research on the evolution of resistance against parasites in crop plants, reciprocal evolution between species in natural plant populations, and mathematical models of the dynamics of coevolution. Recent studies have begun to challenge the prevailing view that natural selection within local plant populations is the major evolutionary process driving this form of coevolution. The emerging pattern from these studies suggests that metapopulation structure, including the effects of gene flow and genetic drift, may be at least as important as local natural selection in determining the genetic dynamics and outcomes of these evolutionary arms races.

THERE are at least half a dozen proposed forms of coevolution between species, some involving reciprocal adaptation and others a combination of adaptation and speciation¹⁻³. Two of these forms of coevolution have had a particularly large effect on the study of the evolution of interactions between parasites and hosts. One is Ehrlich and Raven's⁴ hypothesis of how the evolution of defence and counterdefence in interactions between parasites and hosts may lead to the radiation of species through a process of escape-and-radiation coevolution. The other is the concept of gene-for-gene (GFG) coevolution. Flor's⁵⁻⁸ studies of GFG interactions in cultivated flax and Haldane's⁹ view that many polymorphisms are maintained by interactions with parasites have had lasting effects on studies of the importance of major genes and frequency-dependent selection in coevolution. In fact, the first mathematical model of coevolution was explicitly based on the assumption of a GFG interaction¹⁰.

Discussions of the role of coevolution in the evolution of resistance, the maintenance of polymorphisms, the maintenance of recombination and sex, and the dynamics of evolutionary arms races have all often explicitly or implicitly assumed something close to a GFG relationship among the interacting species. The GFG view has also had a major influence on applied biology. It is one of the few directly evolutionary views that has become a major model for research programmes in agriculture. Much of the work on breeding for resistance to plant pathogens in agricultural crops has assumed that virulence in fungal pathogens and resistance in plants are governed by GFG interactions.

The GFG concept has become a useful tool in both evolutionary biology and plant breeding because it suggests very specific ways in which pairs of species interact. It provides the simplest possible view of the genetics of interspecific interactions against which can be tested the actual genetic relationships between species. Increasingly, however, GFG coevolution is in danger of becoming less predictive and falsifiable as the phrase 'gene-for-gene' is applied to an ever broader array of genetic interactions among species. There is now even discussion of 'polygenic' GFG interactions¹¹. Moreover, as different aspects of the ecology of species have been incorporated into recent empirical and mathematical studies of GFG interactions, the original hypothesis is becoming expanded into a family of views that need to be stated explicitly and tested separately if we are to understand the coevolutionary consequences of a GFG form of interaction.

Here we examine recent conclusions from studies of gene-for-gene interactions in agriculture, natural plant populations, and mathematical theory with a view toward generating a better integration of ecology and evolutionary genetics. We suggest a hierarchical view of GFG coevolution that incorporates different

TABLE 1 The expected compatibility between homozygous genotypes in a single-locus gene-for-gene interaction

Pathogen genotype	Host genotype	
	<i>RR</i>	<i>rr</i>
<i>VV</i>	Incompatible	Compatible
<i>vv</i>	Compatible	Compatible

R is a dominant host gene conferring resistance to the pathogen and *r* is a recessive host gene conferring susceptibility. *V* is a dominant pathogen gene conferring avirulence and *v* is a recessive pathogen gene conferring virulence.

degrees of complexity in population structure and different evolutionary processes as the determinants of patterns in these interactions.

Criteria

The GFG hypothesis as originally formulated by Flor⁷ was broadly cast, stating that "for each gene that conditions reaction in the host there is a corresponding gene that conditions pathogenicity in the parasite". The host and parasite genes have since become known as corresponding genes, and the GFG hypothesis is now generally accepted to mean that "for each gene determining resistance in the host there is a corresponding gene for avirulence in the parasite with which it specifically interacts"¹². According to this hypothesis, the occurrence of a resistant or incompatible reaction depends on both the presence of a gene for resistance (*R*) in the host and the corresponding gene for avirulence (*V*) in the parasite. If a host lacks a specific resistance gene, the corresponding avirulence gene in the parasite cannot be detected. Similarly, if the parasite lacks a specific avirulence gene, the corresponding resistance gene in the host cannot be detected (Table 1).

In diploid host-parasite associations, the interactions occurring at any particular corresponding pair of loci may be further complicated by the occurrence of resistance or avirulence loci in a heterozygous state. This generates a further five genetic combinations (*Vv RR*, *Vv Rr*, *Vv rr*, *VV Rr* and *vv Rr*). Nevertheless, because resistance is usually dominant to susceptibility and avirulence dominant to virulence, these combinations are generally phenotypically indistinguishable from other compatible or incompatible reactions.

The specificity of these interactions, changing from compatible to incompatible as the same host is challenged by different parasite genotypes or as different host genotypes are challenged by the same parasite genotype (Table 1), is the most basic ingredient of a GFG interaction. Consistent congruence of phenotypic responses with the theoretical expectations is taken

as evidence of a GFG relationship. Often a GFG interaction is inferred from the specificity of host-parasite reactions in the matrix of compatibility reactions coupled with an analysis of the genetic basis of resistance. Such genetic analyses, however, have actually been done for only a very small number of associations and often for only a few resistance and avirulence genes. Even such limited analyses, however, are often extremely difficult, if not impossible, to do. In a surprisingly large number of associations, for example, one or both species (although usually the parasite) seldom or never undergoes meiosis. Consequently, most genetic evidence for GFG interactions comes from a small number of well studied cases.

Plant breeding

Until recently, the only evidence for GFG interactions came from studies of crop plants. One of the short-term practical consequences of GFG interactions is that plant breeders have been able to counter virulence in some parasites with single genes for resistance in plants. Consequently, for most of this century this observation has implicitly or explicitly served as the guiding principle in the breeding of crops for resistance to particular fungal pathogens. Since the recognition of the dominant, single-gene nature of resistance against some fungal pathogens, breeding in many crops has involved sequential release of new varieties carrying different resistance genes as a form of evolutionary arms race. This process has resulted in a series of now classic examples of microevolution as pathogens have responded to these new selective forces by acquiring the corresponding genes for virulence. The best examples of these reciprocal responses are undoubtedly found among the cereal rusts (*Puccinia* spp.), where the sequential deployment of single-gene resistance has led to changes in both the frequency of particular multilocus virulence phenotypes (races) in the pathogen or the appearance of totally new races.

Gene-for-gene interactions involving the formal analysis of at least a few genes for resistance and avirulence have now been determined in at least 12 agricultural plant-pathogen associations (Box). In all of these associations, resistance and avirulence are largely inherited as single genes although additional complications have occasionally been detected (for example, expression of resistance or avirulence controlled by two genes or influenced by modifier genes). Furthermore, although both resistance genes in hosts and avirulence genes in parasites commonly segregate independently, this is not always so. Tight linkage groups have been detected among both resistance and virulence genes¹³. Using the less rigorous criteria provided by phenotypic patterns in the host-pathogen matrix plus studies of host genetics alone, GFG interactions have also been postulated to exist in at least 29 other interactions between plants and pathogens, including fungi, viruses and bacteria (Box).

This group of more than 40 associations encompasses a range of host and pathogen genera and a broad diversity of plant life forms including grasses, annual and perennial herbs, shrubs and trees (Box). Most GFG examples involve associations in which the pathogen is highly specialized (so-called biotrophic pathogens like rusts, smuts and mildews), but in some others the pathogens kill host tissue directly (the necrotrophic pathogens). This occurrence of GFG interaction in such a wide variety of pathogens and their host plants supports the idea that this kind of interaction is likely to be found in many host-pathogen associations¹⁴.

Although the GFG model has been a powerful tool in plant breeding for fungal pathogens, it is not clear to what extent such interactions occur in other relationships between plants and their enemies. GFG relationships have been postulated for interactions between plants and some phytophagous insects and nematodes¹⁵⁻¹⁸, but in most cases even less is known about the formal genetics of these associations than about fungal pathogens. The most well known GFG research programme

BOX 1 Demonstrated (*) or suggested gene-for-gene interactions between plants and pathogens

Host and pathogen	Reference [†]
Fungal pathogens	
Basidiomycetes Uredinales (rusts)	
<i>Avena</i> - <i>Puccinia graminis</i> *	27
- <i>Puccinia coronata</i> *	56
<i>Coffea</i> - <i>Hemileia vastatrix</i>	27
<i>Glycine</i> - <i>Phakopsora pachyrhizi</i>	27
<i>Helianthus</i> - <i>Puccinia helianthi</i>	27
<i>Linum</i> - <i>Melampsora lini</i> *	27
<i>Pinus</i> - <i>Cronartium quercuum</i>	57
<i>Triticum</i> - <i>Puccinia graminis</i> *	27
- <i>Puccinia recondita</i> *	27
- <i>Puccinia striiformis</i>	27
Basidiomycetes Ustilaginales (bunts and smuts)	
<i>Avena</i> - <i>Ustilago avenae</i> *	27
<i>Hordeum</i> - <i>Ustilago hordei</i> *	27
<i>Triticum</i> - <i>Ustilago tritici</i>	27
- <i>Tilletia caries</i>	27
- <i>Tilletia controversa</i>	27
Ascomycetes Erysiphales (powdery mildews)	
<i>Hordeum</i> - <i>Erysiphe graminis</i> *	27
<i>Secale</i> - <i>Erysiphe graminis</i>	58
<i>Senecio</i> - <i>Erysiphe fischeri</i>	29
<i>Triticum</i> - <i>Erysiphe graminis</i> *	27
Oomycetes Peronosporales	
Peronosporaceae (downy mildews)	
<i>Lactuca</i> - <i>Bremia lactucae</i> *	27
<i>Medicago</i> - <i>Peronospora trifoliorum</i>	59
<i>Sorghum</i> - <i>Peronosclerospora sorghi</i>	60
Pythiaceae	
<i>Glycine</i> - <i>Phytophthora megasperma</i>	61
<i>Solanum</i> - <i>Phytophthora infestans</i>	27
Other fungi	
<i>Brassica</i> - <i>Plasmodiophora brassicae</i>	62
<i>Callistephus</i> - <i>Fusarium oxysporum</i>	63
<i>Hordeum</i> - <i>Rhynchosporium secalis</i>	64
<i>Malus</i> - <i>Venturia inaequalis</i> *	27
<i>Oryza</i> - <i>Piricularia oryzae</i>	27
<i>Phaseolus</i> - <i>Colletotrichum lindmuthianum</i>	27
<i>Saccharum</i> - <i>Bipolaris sacchari</i>	65
<i>Solanum</i> - <i>Synchytrium endobioticum</i>	27
Bacteria	
<i>Capsicum</i> - <i>Xanthomonas campestris</i>	66
<i>Gossypium</i> - <i>Xanthomonas malvacearum</i> *‡	67
<i>Pisum</i> - <i>Pseudomonas syringae</i>	68
Viruses	
<i>Capsicum</i> -tobacco mosaic	69
<i>Glycine</i> -soybean mosaic	70
<i>Lycopersicon</i> -tobacco mosaic	27
-spotted wilt	27
<i>Phaseolus</i> -bean yellow mosaic	71
<i>Solanum</i> -potato virus x	27

† Complete references listed for most studies before 1986 are given in ref. 27. Complete references for recent studies, plus several older studies not listed ref. 27, are given in the references section of this paper. Citations are to the original studies suggesting or demonstrating a gene-for-gene interaction.

‡ Gene-for-gene relationship demonstrated through comparison of known resistance genes with cloned avirulence genes in the bacterium.

between plants and insects is that between wheat and Hessian fly. Twenty genes for resistance to Hessian fly are now known in wheat, and there are 11 fly biotypes that differ in their abilities to attack wheat varieties with these various resistance genes¹⁹. Diehl and Bush²⁰, however, cautioned that the GFG view of this interaction may be partially an artefact of the experimental protocols used. Virulent fly biotypes have usually been obtained by 'purifying' strains over several generations on wheat varieties with different resistance genes, and then using a hybridization test in the F₄ generation to eliminate any heterozygotes. This process is likely to magnify single gene effects.

In some other studies of GFG interactions between insects and plants, insect populations have been divided into biotypes based upon the idea that each biotype, although perhaps a complex of genotypes, may differ monogenically at least for virulence genes. As some of these insects, however, have been studied in more detail, the genetics of virulence have turned out not to be monogenic. For example, the interaction between rice and the rice brown planthopper (*Nilaparvata lugens*) was previously assumed to be a GFG interaction, and the planthoppers were divided into biotypes accordingly. The 'biotypes' are now known, however, to be polygenically determined. Each is a collection of extreme genotypes for virulence selected by particular rice varieties from a more continuous range of virulence^{21,22}. Similarly, greenbug (*Schizaphis graminum*) biotypes are complexes of genotypes in which virulence for wheat is determined by the interaction of three loci¹¹. Moreover, oviposition preference in at least some phytophagous insects is now known to be controlled by different genes than those that control larval performance, making the genetics of insect attack in these interactions potentially more complex than the genetics of attack by pathogens^{23,24}. Because of these life-history differences between insects and pathogens, GFG relationships may in fact be much more common in interactions involving plant pathogens rather than phytophagous insects, but the genetic studies are still too few to know.

Although breeding programmes and models of GFG interactions in agriculture often focus on how natural selection will favour increased virulence in local pathogen or insect populations, recent studies have suggested that the dynamics and evolution of these interactions is sometimes governed by gene flow over long distances. The use of genetic fingerprinting has confirmed that some epidemics in crops are caused by the spread of one or two clones over large areas. Brown *et al.*²⁵ found that the epidemic of powdery mildew on barley in Britain in the mid 1980s was initiated largely by two clones able to overcome barley cultivars carrying the resistance allele *Mla13*. These results, then, argue for a broader ecological and geographical view of the evolution of GFG interactions in some crops rather than a view based on minimizing the rate of natural selection for virulence in local pathogen populations.

Natural populations

Studies of the genetic basis of host-pathogen interactions in natural populations fall into two general approaches: (1) those in which the role of genetic variation is examined through comparisons of host family lines grown in natural situations, and (2) those involving detailed studies of individual host $\times$ parasite isolates and formal genetic analysis of resistance. The first compares the total resistance phenotype expressed by individuals or family lines against what potentially may be a genetically variable parasite population. As a consequence, the action of major genes for resistance in the host, and avirulence in the parasite, is likely to be masked and confounded by the interplay of gross-morphological attributes (for example, earliness) with quantitative and qualitatively based resistance factors.

The alternative approach of studying individual host $\times$ parasite isolates has come from increasing interest in the dynamics of individual types of resistance. It has been known for some time that many of the elements of classic GFG interactions, such as resistance conferred by single genes with major phenotypic effects and tightly coupled interactions between host lines and specific pathogen genotypes, are found in a wide array of natural host-pathogen associations²⁶⁻³⁰. Many of these studies emphasize the assessment of resistance expressed towards a few diverse pathogen genotypes by individual host lines collected from a range of populations³¹. For this reason, these studies are largely divorced from the ecological context of natural associations and hence provide little information on how GFG interactions are influenced by population structure.

At the other extreme, studies that concentrate heavily on interactions in single populations tend to give the impression of tightly coevolved GFG associations between host and pathogen at the level of the individual population.

Recently, simultaneous study of variation in hosts and pathogens in a number of populations of *Linum marginale* and *Melampsora lini* growing in the same geographic region have provided a new perspective in our understanding of how plants and pathogens continue to coevolve in natural systems³²⁻³⁴. This association is an endemic Australian version of the interaction used by Flor⁵ to first elucidate the GFG concept and provides further evidence that GFG interactions do occur in natural populations. More importantly, however, the study has clearly indicated the spatial scale over which such coevolving associations are likely to develop³⁵. In individual populations, the genetic structure of pathogen demes (that is, the relative frequencies of different races within local pathogen populations) shows little correlation with that of co-occurring hosts³⁴. Instead, stochastic fluctuations during periodic population crashes appear to be a prime cause of marked year-to-year variation in the size and genetic constitution of pathogen demes³³. At its extreme this process may result in the extirpation of local populations. Such fluctuations act independently in different pathogen demes and, with no gene flow, would ultimately lead to substantial genetic differentiation between adjacent populations. In the *Linum-Melampsora* interaction, however, migration occurs at a sufficiently high rate as to ensure interchange between individual populations and the re-establishment of those genes lost through local extinction.

In this perennial host system, changes in the plant are inevitably slower than in the pathogen. Nevertheless, within individual populations the virulence structure of the associated pathogen deme is such that all host individuals are vulnerable to attack regardless of the resistance genes carried by any individual. Changes in the relative frequency of individual resistant lines appear to show little correlation with the frequency of pathogen races in the immediately associated pathogen deme. Thus at the spatial scale of the individual host and pathogen demes there is no clear pattern of reciprocal responses in the genetic structure of the interacting populations. Rather, individual populations are strongly influenced by genetic drift, extinction and gene flow that occur at random among demes within the same epidemiological region. The sum of these processes across many individual host and pathogen demes constitutes the metapopulation level at which GFG coevolution takes place and at which genetic variation in resistance and virulence genes is maintained.

Mathematical models

Although agricultural and coevolutionary GFG models for natural populations differ in their assumptions and goals, recent mathematical models incorporating more aspects of the ecology of species than older models show a tendency toward irregular, and sometimes chaotic, fluctuations in allele frequencies. Agricultural models of GFG interactions have, as their primary objective, prediction of the rate at which a pest population will overcome a particular strategy of deployment of major resistance genes in the plants. For example, Cox and Hatchett's¹⁷ and Gould's³⁶ models of the interaction between Hessian fly and wheat were designed to predict how long it would take a Hessian fly population to reach fixation for virulence genes that overcome corresponding resistance genes deployed either sequentially, in mixtures of cultivars with different resistance genes, or as cultivars carrying multiple resistance genes (pyramiding). Similar models have been constructed for interactions between plants and pathogens³⁷. The models for both insects and pathogens show that, besides the deployment strategy of resistance genes, the mode of reproduction in the pests, their dispersal rate and, for insects, sexual selection all have large effects on the rate at which virulence increases in these populations.

In most GFG models of natural populations, both plants and pathogens reproduce naturally and the objective is to understand the dynamics of gene frequencies and the path of coevolution. Most of the strictly genetic models of GFG coevolution—that is, those that do not incorporate demographic structure and population dynamics—show cyclical fluctuations in gene frequencies, although these fluctuations may be chaotic³⁸. Costs of resistance and virulence, and time delays in frequency-dependent selection, are often important variables determining if cycling in gene frequencies is maintained within these models³⁹. These time delays can result in 'boom-and-bust' cycles, as plant and pathogen numbers cycle slightly out of phase with each other. In Barrett's⁴⁰ models, asymmetries in the relative costs of resistance and virulence can lead to rapid fixation of parasite genotypes and coevolutionary arms races dependent on the continued evolution of novel host genotypes rather than to frequency-dependent selection and polymorphism. Some other recent versions of genetic models, incorporating at least two loci and resulting in three or more genotypes within a population, produce cycles that are even more complex and chaotic, and sometimes qualitatively different, from one-locus models⁴¹.

Most models of the evolution of sex that are driven by coevolution between parasites and hosts are specific forms of GFG coevolution. For example, among recent analyses the models of May and Anderson⁴², Bell and Maynard Smith³⁹, and Hamilton *et al.*⁴³ are based on matching genes for resistance in hosts and virulence in parasites as the genetic basis for evolutionary arms races. In the models of Hamilton *et al.*⁴³, which also consider how a GFG interaction between two species is affected by the presence of additional parasites, the genetic outcome of these interactions is influenced especially by the number of parasites attacking the host and the number of loci involved in defence against parasites. Together, models for the evolution of sex driven by parasite attack increasingly emphasize the importance of ecological factors in driving GFG coevolution.

When other ecological factors, such as changes in the abundance of host populations or different transmission rates, are built into GFG models, a whole range of outcomes becomes possible, and chaotic fluctuations are a common result^{44,45}. The models of Anderson and May^{42,44,46,47} combine epidemiology with genetics and show that epidemiological factors such as density dependence produce polymorphisms that are more likely to be cyclic or chaotic than purely genetic models. The biology and ecology of the interaction are major determinants of how chaotic the fluctuations become⁴⁸. In a recent analysis of a haploid model allowing for fluctuating numbers of hosts and pathogens, Frank⁴⁵ found that maintenance of polymorphism can depend more on ecological and demographic factors such as birth and death rates than on genetic factors such as the pleiotropic costs of different resistance alleles or virulence alleles.

Evolutionary processes

The increasing frequency of chaotic fluctuations in models of GFG, and the geographic aspects of GFG interactions found in recent studies in both agriculture and natural populations, provoke the question of whether GFG interactions within single, natural populations ever produce stable polymorphisms maintained by frequency-dependent selection or a combination of frequency-dependent and density-dependent selection. The problem may be that current models, although increasingly complex, still lack biological realism. No model of single populations, for example, simultaneously includes diploid hosts and parasites (to explore the consequences of dominance), multiple loci, sexual reproduction, and fluctuating population sizes. Of course, this is a never-ending process, and models will always be criticized for being over-simplified. The important point is that as new ecological variables have been added to GFG models, fluctuations in allele frequencies have commonly become increasingly irregular.

Frequency-dependent polymorphisms, whether stable or transient, may well be a common result of interactions between parasites and hosts. Polymorphisms, however, may just as likely result from any of a group of other causes including metapopulation structure (that is, groups of local populations linked by gene flow) with moderate levels of gene flow among demes and interactions of hosts with multiple species of parasites. Few models have explored in detail the effects of these kinds of influences on GFG coevolution, although some have examined epidemiological aspects of population structure⁴⁴ or some aspects of subdivision of populations⁴⁹. A recent model by Hassell, Comins and May⁵⁰, however, provides some clues to how metapopulation structure may influence the maintenance and dynamics of polymorphisms. Although their model was constructed specifically to evaluate the population dynamics of insects and their parasitoids rather than the genetic dynamics of plants and their parasites, the results indicate that metapopulation structure may allow for the regional persistence of interactions that are locally unstable. Small changes in reproductive rate in the host and in gene flow among host and parasitoid subpopulations can lead to very different patterns in the spatial dynamics of these interactions. Hence the metapopulation structure of interacting populations may potentially magnify or mitigate the 'boom-and-bust' cycles expected in some genetic models of pathogen-plant interactions.

Hierarchy

The recent combination of results from agriculture, natural populations and mathematical models suggests that the GFG view of coevolution is developing into a family of hypotheses that differ in their assumptions and predictions. We suggest below a hierarchy of views based on complexity of population structure and the evolutionary processes driving genetic changes in plants and their parasites.

Local agricultural gene-for-gene coevolution. GFG interactions as used in agriculture generate a form of human-induced coevolution that incorporates the simplest forms of host population structure: genetic uniformity or a mixture of a small number of genotypes homozygous for particular resistance alleles. Gene frequencies in the host population are determined completely by planting regimes, and allele frequencies in local populations may swing from zero to a hundred per cent among years. Consequently, the major problem to solve at the local level is how to construct crop populations in ways that minimize the rate at which parasites overcome resistance in the crops.

Geographic agricultural gene-for-gene coevolution. The long-term fate of major genes for resistance in some crop plants appears at least sometimes to be governed as much by geographic events as by local selection on pest populations. Major epidemics of fungal and insect pests have been caused by single races that have spread over large geographic areas, taking advantage of large monocultures of particular plant genotypes. Breeding for durable resistance in crops based upon a GFG model may therefore require different strategies when changes in the genetic composition of the parasite population is governed as much by broad geographic events as it is by selection on a local population.

Panmictic gene-for-gene coevolution. Most conceptual and mathematical work on GFG coevolution in natural populations is based on the view that coevolution occurs at the level of single, large populations of hosts and parasites. Unlike either local or geographic agricultural GFG coevolution, both host and parasite populations reproduce naturally and are genetically variable at the local level. The major assumptions of panmictic GFG coevolution with respect to population structure are that no genetic drift occurs within populations and no gene flow occurs between neighbouring populations. Consequently, the only evolutionary processes that guide coevolution are the mutation rate and the efficacy of natural selection. Frequency-dependent selection and, depending on the mode of transmission,

density-dependent selection become the modes of selection that drive coevolution of the populations, leading to stable or unstable polymorphisms.

Metapopulation gene-for-gene coevolution. If populations of the host or parasite are subdivided into separate demes, then the evolutionary processes driving coevolution may be quite different from those in panmictic models. If individual demes are at least moderately large and gene flow is great enough, then the various demes will effectively act as one large, panmictic population as shown in standard population genetic models. If, however, demes are smaller, population sizes vary greatly over time, and gene flow occurs but to a more limited extent, then mutation rates and frequency-dependent selection within populations may not be the most important evolutionary processes driving GFG coevolution. Frequency-dependent selection may occur within demes, but it may not be responsible for the overall pattern of change in gene frequencies over time. Instead, coevolution may be a regional process governed by a combination of gene flow, genetic drift, and various forms of natural selection.

Continued progress in breeding for durable resistance in plant populations and in understanding the evolutionary dynamics of GFG coevolution in natural populations will require studies that carefully distinguish between these different views of the relative importance of local and geographic processes.

Polygenic inheritance

Some researchers have argued that GFG interactions are artefacts of plant breeding and may have only limited application to natural populations. Sidhu⁵¹, for example, argued that successful application of the GFG concept in agriculture occurs under conditions in which a crop is subject to attack mostly by one particular parasite or pathogen species. Under more natural conditions in which a host is regularly subject to attack by multiple parasite species, evolution of resistance is more likely to be polygenic or at least give the appearance of being polygenic because the number of resistance loci is greater and there may be epistatic (that is, synergistic) interactions among them. Similarly, Barrett⁵² argued that the process of selection of extreme genotypes during plant breeding magnifies the

impression that single-gene effects are common.

It is certainly easy enough to find examples of resistance spanning the entire spectrum from major single gene effects to additivity across several to many loci^{13,27,53,54}. Single crops such as wheat and maize include examples of resistance ranging from monogenic through oligogenic to polygenic⁵⁵. Moreover, the ability of some parasites to survive and develop on plants is known to be determined by at least several genes²⁴. The problem, however, is not simply to understand if an interaction is governed by GFG interactions or more complicated polygenic interactions. Nor should the suggestion that GFG interactions are artefacts of particular breeding and ecological conditions be viewed as a cause for abandonment of the GFG concept. Rather these concerns should be used as an opportunity to ask where and how natural selection may favour coevolution governed by GFG interactions rather than by polygenic effects. Probably all interactions between species that involve GFG relationships also involve polygenic relationships as well. The different conditions of agriculture and natural populations provide the opportunity to ask how different aspects of population structure magnify or mitigate GFG effects.

The emerging pattern from agriculture, natural population and mathematical models is that metapopulation structure is an important influence on how interactions evolve. The geographic structure of populations may, along with its other effects on the dynamics of coevolution, influence where and when major genes rather than polygenes are favoured in the evolutionary arms races between species. Studying the role of metapopulation structure in shaping the genetics of coevolution is obviously not going to be an easy task. But only by linking our developing understanding of the geographic structure of populations with analyses of GFG and other forms of coevolution will it be possible to make progress in understanding the role that coevolution plays in structuring populations and shaping the evolution of interactions. □

John N. Thompson is at the Departments of Botany and Zoology, Washington State University, Pullman, Washington 99164, USA; Jeremy J. Burdon is at the Division of Plant Industry, CSIRO, GPO 1600, Canberra 2601, Australian Capital Territory, Australia.

- Thompson, J. N. *Interaction and Coevolution* (Wiley, New York, 1982).
- Thompson, J. N. *Trends Ecol. Evol.* **4**, 179–183 (1989).
- Thompson, J. N. in *Pests, Pathogens and Plant Communities* (eds Burdon, J. J. & Leather, S. R.) 249–271 (Blackwell, Oxford, 1990).
- Ehrlich, P. R. & Raven, P. H. *Evolution* **18**, 586–608 (1964).
- Flor, H. H. *Phytopathology* **32**, 653–669 (1942).
- Flor, H. H. *Phytopathology* **43**, 624–628 (1953).
- Flor, H. H. *Adv. Genet.* **8**, 29–54 (1956).
- Flor, H. H. *Rev. Phytopath.* **9**, 275–296 (1971).
- Haldane, J. B. S. *La Ricerca Scien.* **19** (suppl.), 68–76 (1949).
- Mode, D. J. *Evolution* **12**, 158–165 (1958).
- Puterka, G. J. & Peters, D. C. *Genome* **32**, 109–114 (1989).
- Kerr, A. A. *Rev. Phytopath.* **25**, 87–110 (1987).
- Christ, B. J., Person, C. O. & Pope, D. D. in *Populations of Plant Pathogens: their Dynamics and Genetics* (eds Wolfe, M. S. & Caten, C. E.) 7–19 (Blackwell, Oxford, 1987).
- Person, C. O. *Can. J. Bot.* **37**, 1101–1130 (1959).
- Parrott, D. M. *Nematologica* **27**, 372–384 (1981).
- Sidhu, G. S. *Bot. Rev.* **47**, 387–419 (1981).
- Cox, T. S. & Hatchett, J. H. *Environ. Ent.* **15**, 24–31 (1986).
- Saxena, R. C. & Barrion, A. A. *Insect Sci. Applic.* **8**, 453–458 (1987).
- Weller, S. J., Ohm, H. W., Petterson, F. L., Foster, J. E. & Taylor, P. L. *Crop Sci.* **31**, 1163–1168 (1991).
- Diehl, D. R. & Bush, G. L. A. *Rev. Entomol.* **29**, 471–504 (1984).
- Den Hollander, J. & Pathak, P. K. *Ent. Exp. Appl.* **29**, 76–86 (1981).
- Claridge, M. F. & Den Hollander, J. *Ent. Exp. Appl.* **32**, 213–221 (1982).
- Thompson, J. N. *Evolution* **42**, 1123–1234 (1988).
- Thompson, J. N., Wehling, W. & Posolsky, R. *Nature* **344**, 148–150 (1990).
- Brown, J. K. M., Jessop, A. C. & Rezanoor, H. N. *Proc. R. Soc. B246*, 83–90 (1991).
- Parker, M. A. *Evolution* **39**, 713–723 (1985).
- Burdon, J. J. *Diseases and Plant Population Biology* (Cambridge Univ. Press, Cambridge, 1987).
- Burdon, J. J. *Oecologia* **73**, 257–267 (1987).
- Harry, I. B. & Clarke, D. D. *New Phytol.* **107**, 715–723 (1987).
- Jarosoz, A. M. & Burdon, J. J. *Heredity* **64**, 347–353 (1990).
- Moseman, J. G., Nevo, E., El Morshidy, M. A. & Zohary, D. *Euphytica* **33**, 44–47 (1984).
- Burdon, J. J. & Jarosz, A. M. *Evolution* **45**, 205–217 (1991).
- Burdon, J. J. & Jarosz, A. M. *Plant Path.* **41**, 165–179 (1992).
- Jarosoz, A. M. & Burdon, J. J. *Evolution* **45**, 1618–1627 (1991).
- Burdon, J. J., Brown, A. H. D. & Jarosz, A. M. in *Pests, Pathogens and Plant Communities* (eds Burdon, J. J. & Leather, S. R.) 233–247 (Blackwell, Oxford, 1990).
- Gould, F. *Environ. Ent.* **15**, 1–10 (1986).
- Kiyosawa, S. A. *Rev. Phytopath.* **20**, 93–117 (1982).
- Levin, S. A. in *Coevolution* (ed. Nitecki, M.) 21–65 (University of Chicago Press, Chicago, 1983).
- Bell, G. & Maynard Smith, J. *Nature* **328**, 66–68 (1987).
- Barrett, J. A. *Phil. Trans. R. Soc. B319*, 473–483 (1988).
- Seger, J. *Phil. Trans. R. Soc. B319*, 541–555 (1988).
- May, R. M. & Anderson, R. M. *Proc. R. Soc. B219*, 281–313 (1983).
- Hamilton, W. D., Axelrod, R. & Tanese, R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3566–3573 (1990).
- Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, Oxford, 1991).
- Frank, S. A. *Heredity* **67**, 73–83 (1991).
- Anderson, R. M. & May, R. M. *Parasitology* **85**, 411–426 (1982).
- May, R. M. in *Ecology and Genetics of Host-Parasite Interactions* (eds Rollinson, D. & Anderson, R. M.) 243–262 (Academic, London, 1985).
- May, R. M. & Anderson, R. M. *Parasitology* **100**, S89–101 (1990).
- Frank, S. A. *Trends Genet.* **8**, 213–219 (1992).
- Hassell, M. P., Comins, H. N. & May, R. M. *Nature* **353**, 255–258 (1991).
- Sidhu, G. S. *Bioscience* **34**, 368–373 (1984).
- Barrett, J. A. in *Ecology and Genetics of Host-Parasite Interactions* (eds Rollinson, D. & Anderson, R. M.) 215–225 (Academic, New York, 1985).
- Osborn, T. C., Alexander, D. C., Sun, S. S. M., Cardona, C. & Bliss, F. A. *Science* **240**, 207–210 (1988).
- Simms, E. L. & Rausher, M. D. *Evolution* **43**, 573–585 (1989).
- Khush, G. S. & Brar, D. S. *Adv. Agron.* **45**, 223–274 (1991).
- Nof, E. & Dinor, A. *Phytoparasitica* **9**, 240 (1981).
- Kinloch, B. B. Jr. & Walkinshaw, C. H. in *Rusts of Pine* (eds Hiratsuka, Y., Samoil, J. K., Blenis, P. V., Crane, P. E. & Laishley, B. L.) 219–228. (*For. Canada Northwest Region Inf. Rep. NOR-X-317*, 1989).
- Geiger, H. H., Schuhmacher, A. E. & Billenkamp, N. *Plant Breeding* **100**, 97–103 (1988).
- Skinner, D. Z. & Stuteville, D. L. *Phytopathology* **75**, 119–121 (1985).
- Pawar, M. N., Frederiksen, R. A., Mughogho, L. K. & Bonde, M. R. *Phytopathology* **75**, 1374 (1985).
- Layton, A. C. & Kuhn, D. N. *Phytopathology* **78**, 961–966 (1988).
- Crute, I. R., Phelps, K., Barnes, A., Buczacki, S. T. & Crisp, P. *Pl. Path.* **32**, 405–420 (1983).
- Persiel, F. & Lein, H. J. *Plant Dis. Protect.* **96**, 47–59 (1989).
- Allard, R. W. J. *Heredity* **81**, 1–6 (1991).
- Kang, M. S. & Dean, J. L. *Plant Disease* **71**, 450–451 (1987).
- Minsavage, G. V. et al. *Molec. Pl. Microbe Interact.* **3**, 41–47 (1990).
- Gabriel, D. W., Burges, G. R. & Lazo, G. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6415–6419 (1986).
- Taylor, J. D., Bervan, J. R., Crute, I. R. & Reader, S. L. *Pl. Path.* **38**, 364–375 (1989).
- Tobias, I., Fraser, R. S. S. & Gerwitz, A. *Physiol. Molec. Pl. Path.* **35**, 271–286 (1989).
- Roane, C. W., Tolin, S. A. & Buss, G. R. *Soybean genet. Newsl.* **13**, 136–139 (1986).
- Schmidt, H. E. & Zobywalski, S. *Arch. Phytopath. Pflanzen* **20**, 95–96 (1984).

ACKNOWLEDGEMENTS. We thank A. H. D. Brown, G. J. Lawrence and G. F. Moran for comments on the manuscript. This work was supported by a Fulbright Senior Fellowship through the Australian-American Exchange Foundation to J.N.T. and by grants from the NSF.

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to presubmission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and should contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the Commentary Editor in the form of a one-page synopsis.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with each copy of a submitted manuscript, and clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.

Structure of a C-type mannose-binding protein complexed with an oligosaccharide

William I. Weis^{*†}, Kurt Drickamer & Wayne A. Hendrickson^{*}

Department of Biochemistry and Molecular Biophysics and ^{*}Howard Hughes Medical Institute, Columbia University, 630 West 168 Street, New York, New York 10032, USA

C-type (Ca^{2+} -dependent) animal lectins such as mannose-binding proteins mediate many cell-surface carbohydrate-recognition events. The crystal structure at 1.7 Å resolution of the carbohydrate-recognition domain of rat mannose-binding protein complexed with an oligomannose asparaginylo-oligosaccharide reveals that Ca^{2+} forms coordination bonds with the carbohydrate ligand. Carbohydrate specificity is determined by a network of coordination and hydrogen bonds that stabilizes the ternary complex of protein, Ca^{2+} and sugar. Two branches of the oligosaccharide crosslink neighbouring carbohydrate-recognition domains in the crystal, enabling multivalent binding to a single oligosaccharide chain to be visualized directly.

THE family of calcium-dependent carbohydrate-binding proteins designated C-type animal lectins¹ includes endocytic receptors of hepatocytes and macrophages²⁻⁴, the selectin cell-adhesion molecules⁵⁻⁸, and secreted molecules found in the extracellular matrix and in serum^{9,10}. Each of these proteins contains one or more carbohydrate-recognition domains (CRDs) combined with domains responsible for the physiological functions of the molecule. Mammalian mannose-binding proteins (MBPs) are serum C-type lectins that function in immunoglobulin-independent host defence against pathogens. These proteins bind to oligomannose cell surface oligosaccharides of various bacteria and fungi, and neutralize them by complement-mediated cell lysis¹¹ or opsonization¹². The CRD of mannose-binding protein is combined with a collagen-like domain which is presumed to mediate activation of the complement cascade. MBPs are found as oligomers of a trimeric building block formed by a collagenous triple-helical structure¹³.

Despite their shared dependence on calcium for carbohydrate binding, various C-type lectins have distinct carbohydrate-recognition specificities. Each C-type CRD is roughly 120 amino acids long, and contains 14 invariant amino acids and another 18 that are conserved in character¹. The three-dimensional structure of the CRD from rat mannose-binding protein A (MBP-A) was recently determined¹⁴ by X-ray diffraction analysis of crystals in which the two required calcium ions (Ca^{2+}) were replaced with holmium ions (Ho^{3+}). Although that structure provides an explanation for the conserved sequence motif and indicates that the basic fold of the MBP CRD is common to the C-type lectins, it does not indicate how these proteins bind to carbohydrates. Here we describe the structure at 1.7 Å resolution of the Ca^{2+} -containing MBP-A CRD complexed with an oligomannose asparaginylo-oligosaccharide. The structure makes clear the basis of the strict Ca^{2+} requirement for carbohydrate binding of C-type lectins, and provides an explanation of the carbohydrate specificity of MBP-A. The binding of a single branched oligosaccharide to two CRDs in the crystal yields insights into the nature of multivalent interactions that produce the high-affinity binding of lectins to carbohydrates at the cell surface.

Crystal structure of the complex

We have previously described¹⁵ crystals of a complex between a dimeric fragment of MBP-A that contains the minimal CRD, and a $\text{Man}_6\text{GlcNAc}_2\text{Asn}$ asparaginylo-oligosaccharide from ovalbumin (Fig. 1a). The structure of the complex was deter-

mined by molecular replacement methods and refined to 1.7 Å resolution using this and a closely related crystal form (Table 1). The conformations of the N-terminal β -strands of the two protomers, which form part of the dimer interface, are identical to one another in the two Ca^{2+} -containing crystal forms, but they differ in the Ho^{3+} -containing crystals¹⁴. Furthermore, the dimer has a different protomer relationship in each of these crystal forms (Table 1). These observations lend support to the notion that a variable dimer interface results from the proteolytic preparation and is not a conserved feature of this domain¹⁴.

A single asparaginylo-oligosaccharide molecule binds to different protomers of neighbouring CRD dimers (Fig. 1) and thus constitutes an integral component of the crystal lattice. The crosslinking of dimers by the oligosaccharide explains why this crystal form does not grow in its absence¹⁵, and the stoichiometry of one glycopeptide per dimer is consistent with the carbohydrate and amino-acid composition obtained from a washed crystal¹⁵. The five mannose residues of the $\alpha(1,2)$ and $\alpha(1,3)$ branches that form the crosslink are well ordered in the refined electron density map, and the glycosidic bond of the $\alpha(1,6)$ branch and the $\beta(1,4)$ bond between mannose and GlcNAc are also visible (Fig. 1b). Although the $\alpha(1,6)$ -branched mannose and the $\text{GlcNAc}_2\text{Asn}$ portions of the asparaginylo-oligosaccharide are not visible and therefore disordered, there is room for these residues in the solvent interstices of the crystal. The structure of a similarly crosslinked complex between wheat-germ agglutinin and an O-linked glycopeptide has been reported recently¹⁶.

In the Ca^{2+} -containing crystal forms of the MBP-A CRD, both protomers contain a third calcium ion site (labelled 3) near calcium 1 (Fig. 2a, b) which is absent in the Ho^{3+} form. The carboxylate groups of Glu 165 and Asp 194 ligate calcium 3, along with four water molecules. Calcium ions 1 and 3 share these two side chains in a symmetrical fashion: both carboxylate oxygen atoms of Glu 165 ligate calcium 1 and one of them is also a ligand of calcium 3, whereas the two carboxylate oxygen atoms of Asp 194 ligate calcium 3 and one is also a ligand of calcium 1. Thus 12 oxygen atoms, 7 from amino acids and 5 from water molecules, provide a full complement of ligands for the 7-coordinate complexation of both calcium ions. Binding data indicate that 2 Ca^{2+} (or Ho^{3+}) bind to each protomer¹⁵, so the third site may arise from adventitious binding of excess Ca^{2+} present in the crystallization medium¹⁵. There are no direct lattice interactions with calcium 3 of either protomer, although both are linked through chains of water molecules to other CRDs in the lattice.

[†] Present address: Department of Cell Biology, Fairchild Building, Stanford University School of Medicine, Stanford, California 94305, USA.

TABLE 1 Molecular replacement, data collection and refinement

Bragg spacing limits (Å)	Independent* reflections	Fraction complete	Fraction >3 σ_1	Multiply measured reflections	$R_{\text{merge}}^\dagger$	Refinement		
						Number used ($ F > 3\sigma_{ F }$)	Fraction of possible	$R_{\text{cryst}}^\ddagger$
10.0–3.36	3,056	0.96	0.99	2,610	0.044	3,042	0.95	0.170
3.36–2.69	2,997	0.96	0.96	2,579	0.063	2,986	0.96	0.162
2.69–2.35	3,025	0.97	0.91	2,499	0.085	2,878	0.92	0.174
2.35–2.14	2,915	0.93	0.87	2,409	0.104	2,811	0.90	0.172
2.14–1.99	2,878	0.92	0.79	2,313	0.130	2,685	0.86	0.169
1.99–1.87	2,907	0.94	0.67	2,226	0.179	2,469	0.80	0.187
1.87–1.78	2,565	0.82	0.50	1,637	0.238	2,033	0.65	0.207
1.78–1.70	2,407	0.75	0.34	1,238	0.291	1,557	0.50	0.238
10.0–1.70	22,750	0.91	0.77	17,511	0.070	20,461	0.82	0.174
Model geometry—r.m.s. deviations from ideality					B value restraints—r.m.s. differences (Å ²)			
Bond lengths		0.009 Å			Main chain bond-related			1.5
Bond angles		1.4°			Main chain angle-related			1.7
Side-chain torsion angles		19.0°			Side chain and carbohydrate bond-related			2.5
Improper torsion angles		1.4°			Side chain and carbohydrate angle-related			2.8
Peptide ω angles		1.9°						

All statistics are from the type II crystal form. **Molecular replacement:** Data from a type I monoclinic crystal (space group $P2_1$; $a=38.6$ Å, $b=85.4$ Å, $c=39.9$ Å, $\beta=91.8^\circ$; 1 dimer/asymmetric unit) were collected as described previously¹⁵. The molecular replacement search model included all protein atoms of the Ho^{3+} -substituted dimer model¹⁴, but excluded Ho^{3+} and water molecules; the refined individual isotropic temperature factors were retained. All calculations were performed with the program X-PLOR³⁶. A rotation function using data between 10 and 4 Å Bragg spacings and a maximum Patterson vector length of 46 Å yielded two solutions related by $\sim 180^\circ$ corresponding to the two possible orientations of the dimer in the asymmetric unit, at 7.8 and 7.5 standard deviations (σ) above the mean value of the rotation function (next peak at 6.8 σ). Because the protomers of the search model are related by an inexact diad axis (176.8°), Patterson-correlation refinement³⁷ of this model was performed first with the dimer as a rigid body, and then with the two protomers as independent rigid bodies. The oriented model was placed in the unit cell with a translation function (10–4 Å), which showed only one significant peak (6.3 σ , next peak 2.0 σ over the mean). Rigid-body refinement of this model gave an R value of 0.346 for all data between 10 and 2.8 Å Bragg spacings. The protomers of this model are related by an exact 180° 2-fold non-crystallographic symmetry axis, oriented as found in a self-rotation function¹⁵, and an $F_o - F_c$ difference Fourier map had strong peaks corresponding to the four expected Ca^{2+} (Ho^{3+}) positions left out of the search model. Inspection of difference Fourier maps revealed that the NH_2 -terminal β -strands are equivalent (see text). The rigid-body refined model was rebuilt accordingly, and a few other minor changes were made in regions of lattice contacts. **Data collection and molecular replacement of type II crystals:** A 1.7 Å data set from a single crystal grown as previously described¹⁵ was collected at 4 °C by rotation through 155° (65 oscillation scans of 2.5–3.0°, 0.2° overlap), with the crystal intentionally set off its major zones to avoid a blind region. Data were recorded on Fuji imaging plates at the F1 beamline of the Cornell High Energy Synchrotron Source ($\lambda=0.90832$ Å; crystal-to-plate distance, 132.75 mm) and digitized with the CHESS reader³⁸. The data from this crystal could not be indexed with the previously reported (type I) lattice; refinement of the unit-cell parameters with an autoindexing algorithm³⁹ revealed that this crystal belonged to a slightly different monoclinic lattice (space group $P2_1$; $a=34.81$ Å, $b=85.16$ Å, $c=39.55$ Å, $\beta=97.41^\circ$; 1 dimer/asymmetric unit), designated type II. Profile-fit Lorentz polarization corrected intensities were obtained using the program DENZO (Z. Otwinowski, unpublished) and reduced with the CCP4 package⁴⁰. Packing considerations and a self-rotation function map indicated that this form is similar enough to the type I crystals to permit direct refinement of the 2.8 Å molecular replacement model into the new lattice. Rigid-body refinement of the type I molecular replacement model, first as a dimer and then as two independent protomers, gave an R value of 0.358 for all data between 10 and 2.8 Å. This solution agrees with a self-rotation function, and was confirmed by repeating the molecular replacement calculations used for the type I form with the new data. **Refinement:** The model of the type II form was refined by conjugate-gradient minimization using X-PLOR. Extension past 2.0 Å Bragg spacings required application of an overall anisotropic temperature factor tensor⁴¹ which was periodically updated during the refinement process (final values: $b_{11}=4.1$ Å²; $b_{12}=b_{23}=0.0$ Å²; $b_{13}=2.2$ Å²; $b_{22}=-4.0$ Å²; $b_{33}=-4.2$ Å²). Empirical energy parameters for amino acids⁴² were used with dihedral and improper torsion angle force constants increased to achieve optimal weighting, as assessed by the free R value⁴³ (W.I.W. and A. T. Brünger, unpublished observations). Carbohydrate parameters were modified from those previously described⁴⁴ to place them on a comparable scale to the protein parameters, and to reflect more accurately bond and angle distributions seen in small molecule crystal structures. Non-crystallographic symmetry restraints were not applied; the main chain and C β atoms of the two protomers superimpose with an r.m.s. deviation of 0.33 Å and a rotation of 179.3° . Residues T177, V199, and N201 (single-letter amino-acid code) of protomer 2, and H116, K152, and D184 of both protomers, lie outside the 'allowed' regions of the Ramachandran diagram (outside the 8 kcal mol⁻¹ contour of the map computed by Peters and Peters⁴⁵); D184 is in a type II turn normally occupied by glycine. A total of 7 amino-acid side chains, one exocyclic mannose hydroxyl group, and one water molecule have been modelled with two conformations. Average isotropic B values are 23.0 Å² for protein and Ca^{2+} , 35.7 Å² for sugars, and 35.1 Å² for water molecules. The final model encompasses residues 109–220 of protomer 1, 109–221 of protomer 2, 6 Ca^{2+} , 5 sugar residues (plus 1 atom each from 2 other sugars), and 211 water molecules. The coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank⁴⁶ (entry 2MSB).

* Measurements were recorded an average of 2.4 times.

$\dagger R_{\text{merge}} = \sum_i \sum_j |I_i(\mathbf{h}) - \langle I(\mathbf{h}) \rangle| / \sum_i \sum_j I_i(\mathbf{h})$ where $I_i(\mathbf{h})$ is the i th measurement and $\langle I(\mathbf{h}) \rangle$ is the weighted mean of all measurements of $I(\mathbf{h})$.

$\ddagger R_{\text{cryst}} = \sum_h |F(\mathbf{h})_{\text{obs}}| - |F(\mathbf{h})_{\text{calc}}| / \sum_h |F(\mathbf{h})_{\text{obs}}|$.

Mannose–CRD interactions

Only the terminal mannose of each visible branch of the oligosaccharide interacts directly with a CRD (Fig. 1b, c); the rest of the carbohydrate projects away from the CRD (Fig. 2a) and makes limited contact with the protein. The key feature of the interaction is the direct ligation of calcium ion 2 by the 3- and 4-hydroxyl groups of a terminal mannose to form an 8-coordinate calcium ion complex (Fig. 2c, d). The coordination complex is a pentagonal bipyramid in which the second apical

position is bisected by the two hydroxyl groups of the sugar. In the sugar-free Ho^{3+} -containing crystal, the cation is 7-coordinate, with a water molecule occupying the position taken up by the mannose 4-OH¹⁴. There are no significant differences in protein structure in this region between these two crystal forms, so the coordination of Ca^{2+} by mannose can be described simply as a replacement of a water molecule by the two sugar hydroxyl groups.

The mannose–CRD complex uses the full complement of

TABLE 2 Comparison of glycosidic bond torsion angles

Linkage	X-ray*			NMR [†]		
	ϕ	ψ	ω	ϕ	ψ	ω
Man 8 $\alpha(1,3)$ Man 5 α	+76	+141	—	+70	+110	—
Man 5 $\alpha(1,6)$ Man 4 β	+67	+177	-66	+60	+160	+60
Man 6 $\alpha(1,3)$ Man 4 β	+78	+116	—	+70	+100	—
Man 9 $\alpha(1,2)$ Man 6 α	+96	+144	—	+75	+100	—
Man 7 $\alpha(1,6)$ Man 5 α	ND	-146	-73	Not reported		
Man 4 $\beta(1,4)$ GlcNAc 3 β	-119	ND	—	Not reported		

Residue numbers are shown in Fig. 1a. Torsion angles are defined by four atoms as follows: $\phi=O5-C1-O1-C'X$, for $X=2, 3, 4, 6$; $\psi=C1-O1-C'X-C'(X+1)$, for $X=2, 3, 4$. For $\alpha(1,6)$ linkages, $\psi=C1-O1-C'6-C'5$, and $\omega=O1-C'6-C'5-O'5$. All angles are given in degrees. ND, not defined in X-ray structure. χ angles ($O5-C5-C6-O6$) for those residues not in $\alpha(1,6)$ linkages are: Man 6, -62° , Man 8, $+50^\circ$ and -75° (equally populated); Man 9, $+65^\circ$.

* Each linkage that is completely defined in the X-ray structure lies in the minimum of the corresponding ϕ , ψ potential surface calculated by empirical⁴⁷ or semi-empirical³² (MNDO) methods, except the $\alpha(1,2)$ linkage, which is in the minimum of the MNDO potential surface, within 3 kcal mol⁻¹ of the minimum of empirical energy potential, and within 2 kcal mol⁻¹ of the NMR-derived conformation. Both $\alpha(1,3)$ linkages are in the same minimum.

[†] From refs 30 and 31; estimated errors of $\pm 15^\circ$ for $\alpha(1,3)$ linkages, $\pm 20^\circ$ for $\alpha(1,2)$ and $\alpha(1,6)$ linkages^{30,48}. Angles have been converted to the crystallographic convention given above.

non-covalent bonding potential of the 3- and 4-hydroxyl groups. One lone electron pair of each hydroxyl forms a coordination bond with the calcium ion, the other lone pair acts as a hydrogen-bond acceptor from an asparagine side chain amide proton, and the hydroxyl proton forms a hydrogen bond with a carboxylate oxygen of glutamic acid (Fig. 2d). The covalent and non-covalent bonds around each hydroxyl oxygen display the tetrahedral arrangement expected for the hydroxyl group. The remaining oxygen of each amino-acid side chain that participates in hydrogen bonds with the sugar serves as a coordination ligand for the calcium ion, thereby producing an intimately linked ternary complex of protein, Ca²⁺ and sugar. The 3- and 4-OH groups of the terminal mannoses participate in the only direct hydrogen-bonds between protein and sugar observed in both protomers.

With one exception, the two visible terminal mannose residues bind identically to the CRD; the amino acids that constitute Ca²⁺ site 2, Ca²⁺ 2, and the terminal mannose bound at each protomer superimpose with an r.m.s. deviation of 0.12 Å. The exocyclic 6-OH of the $\alpha(1,3)$ branch terminus exists in two equally populated conformations (Table 2); one of them ($\chi = -75^\circ$), unique to this branch, can form hydrogen bonds with Glu 193 O ϵ 1 and His 189 N δ 1 (Fig. 1c). The observation of this partially populated conformation in one protomer only suggests that the mannose 6-OH is not a significant determinant of binding specificity, a conclusion reached from binding studies¹⁷ and from consideration of MBP binding specificity (see below).

The mode of binding between mannose and the MBP-A CRD contrasts with that seen in other carbohydrate-binding protein-sugar complexes determined crystallographically. The role of Ca²⁺ as a direct sugar ligand differs fundamentally from the function of metal ions in the legume lectins such as concanavalin A¹⁸, *Erythrina corallodendron* lectin¹⁹, and *Lathyrus ochrus* isolectin I (ref. 20), in which Ca²⁺ and Mn²⁺ maintain structural integrity of the site and help to position amino-acid side chains for binding to sugar, but do not interact directly with the sugar. Also, unlike other lectin-saccharide complexes^{21,22}, comparatively few van der Waals interactions, and no significant aromatic stacking interactions, occur between mannose and the MBP CRD. The bound mannose lies over a portion of the His 189 side chain, but the binding interactions shown in Fig. 2d splay the sugar ring away from the imidazole ring so that although the bound mannose is close enough to the protein to exclude solvent, only the axial 2-OH is in contact with the aromatic ring (3.4 Å from C δ 2) (Figs 1c, 4c). Only two other significant van der Waals contacts occur: mannose C4 is 4.1 Å from His 189 C β (Figs 1c, 4c), and the exocyclic C6 is 4.0 Å from Ile 207

C δ 1 (Figs 1c, 2c) (distances are averaged over the two protomers). Complex formation buries ~ 150 Å² of protein and 250 Å² of carbohydrate solvent-accessible surface area²³ at each protomer; virtually all of the interface arises from the interaction with the terminal mannose (Fig. 1c).

Analysis of single-site mutants in the MBP CRD provides strong support for the model presented here. Changes in amino acids over much of the surface have no effect on carbohydrate binding²⁴, consistent with the limited interaction surface seen in the crystal structure. A particularly informative single amino-acid mutant is Asn 187 $\rightarrow$ Asp. According to the binding scheme shown in Fig. 2d, this change would remove the hydrogen-bond donor of the Asn side chain and replace it with a carboxylate oxygen unable to satisfy the hydrogen-bonding requirements of the mannose 3-OH group. However, the other side chain oxygen would be still present to act as a Ca²⁺ ligand. This mutant would therefore be predicted to bind Ca²⁺, but not sugar. This is the observed phenotype: the mutant displays normal Ca²⁺ affinity, as assessed by resistance to proteolysis (Fig. 3), but is not purifiable on a mannose-affinity column²⁴. Furthermore, the wild-type protein displays a marked enhancement in Ca²⁺ affinity in the presence of mannose (Fig. 3), behaviour that is consistent with an equilibrium in which Ca²⁺ must be bound to the CRD before sugar can bind. In contrast, mannose affects only slightly the Ca²⁺ affinity of the mutant CRD (Fig. 3), confirming the reduced affinity of this protein for sugar.

Specificity of MBP and other C-type lectins

MBPs bind to D-mannose, N-acetyl-D-glucosamine (GlcNAc), D-glucose, and L-fucose, but not to D-galactose and its derivatives. The binding scheme shown in Fig. 2d indicates that the primary determinant of saccharide specificity in MBP is the presence of hydroxyl groups in positions equivalent to the 3- and 4-OH groups of D-mannose. The importance of the 3- and 4-OH groups in the specificity of C-type lectins was anticipated on the basis of binding studies with derivatized sugars¹⁷. N-Acetylglucosamine and glucose differ from mannose in the presence of equatorial rather than axial substituents at the 2-position of the pyranose ring. Superposition of the pyranose rings of these sugars by matching ring positions reveals that the binding scheme seen in the mannose complex can be preserved in the glucose derivatives without steric clashes between the protein and the equatorial-2 substituent (Fig. 4a). The chicken hepatic lectin, which removes from circulation glycoproteins bearing oligosaccharides that contain terminal N-acetylglucosamine², has the same set of calcium-2 ligands as MBP-A, so it would be expected that this receptor recognizes its target glycoproteins in the manner shown in Fig. 4a.

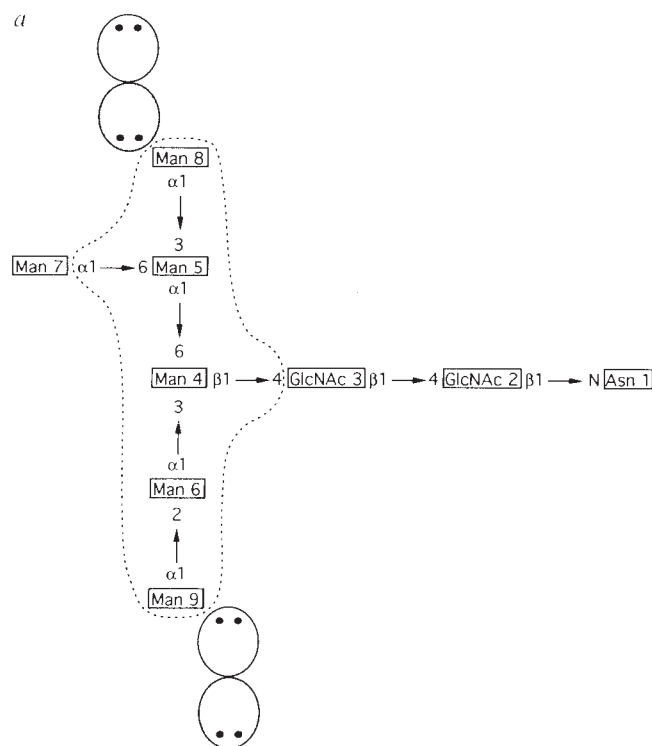


FIG. 1 Oligosaccharide structure and binding. *a*, Structure of the asparaginyl-oligosaccharide used in this study. Dashed line encloses residues and bonds visible in the electron density map. The CRD is represented as an oval, with the two dots denoting the Ca^{2+} -binding end. The carbohydrate links protomer 2 (white) to a different CRD, specifically the $(-x, y + 1/2, 1 - z)$ symmetry mate of protomer 1 (shaded). *b*, Stereo view of the $F_o - F_c$ difference electron density map (2.5σ contour) made from the final model, excluding carbohydrate atoms from the structure factor calculation. The orientation corresponds to that in *a*, with the $\alpha(1, 2)$ branch at the bottom, the $\alpha(1, 3)$ branch at the top, the glycosidic bond to the $\alpha(1, 6)$ branch at the left, and the glycosidic bond to the GlcNAc₂Asn moiety at the right. The crystallographic *b* axis is approximately perpendicular to the page. Carbon atoms of the carbohydrate are shown in yellow, oxygen atoms in red. Ordered water molecules have been omitted for clarity. Protomer 1 is shown in red and protomer 2 in green, and Ca^{2+} 2 appears as a spherical dot surface in each protomer. *c*, Stereo view of the visible oligosaccharide including bound water molecules. The orientation is the same as that in *b*. White, black and grey spheres represent carbon, oxygen, and nitrogen, respectively; black spheres marked 'w' represent water molecules. The large spheres at either end represent Ca^{2+} 2. Hydrogen bonds are shown as dashed lines, and are based on a distance cutoff of 3.3 Å; in some cases potential hydrogen-bonds within this distance have been omitted owing to poor geometry. Amino acids in protomers 1 and 2 are shown with black bonds, and have been labelled with 1,000 and 2,000 added to their sequence numbers, respectively. For clarity, not all atoms of the amino acids have been displayed. The asterisk denotes the side chain of Arg 175 of a third CRD from another unit cell, which is linked to the sugar through two water molecules. The terminus of the protomer-2 Lys 182 side chain appears to be disordered, so its hydrogen bond with the 4-OH of Man 6 in the figure is probably not significant. With this lysine in the position shown, the complex buries an additional 25 Å² of protein and 30 Å² of carbohydrate (from Man 6) solvent-accessible surface over the values cited in the text.

The binding of L-fucose to MBP can also be understood in the context of the binding scheme described for mannose. Although this sugar contains an axial 4-OH, the equatorial 2- and 3-OH groups of its L-pyranose ring can be superimposed on the 3- and 4-OH groups of the D-pyranose ring such that the 4-OH of fucose occupies the same position as the mannose 2-OH, and that the C1 and C6 positions are reversed with respect to mannose²⁵ (Fig. 4*b*). The naturally occurring α -linkages at C1 of fucose can be accommodated on this model with no steric clashes. This observation is noteworthy because the selectins, which bind to carbohydrates containing both terminal fucose and sialic acid^{26–28}, have exactly the same amino-acid residues as MBP in calcium site 2. Therefore, it is likely that calcium site 2 is the fucose-binding site in the selectins, and that fucose binds to these proteins as shown in Fig. 4*b*.

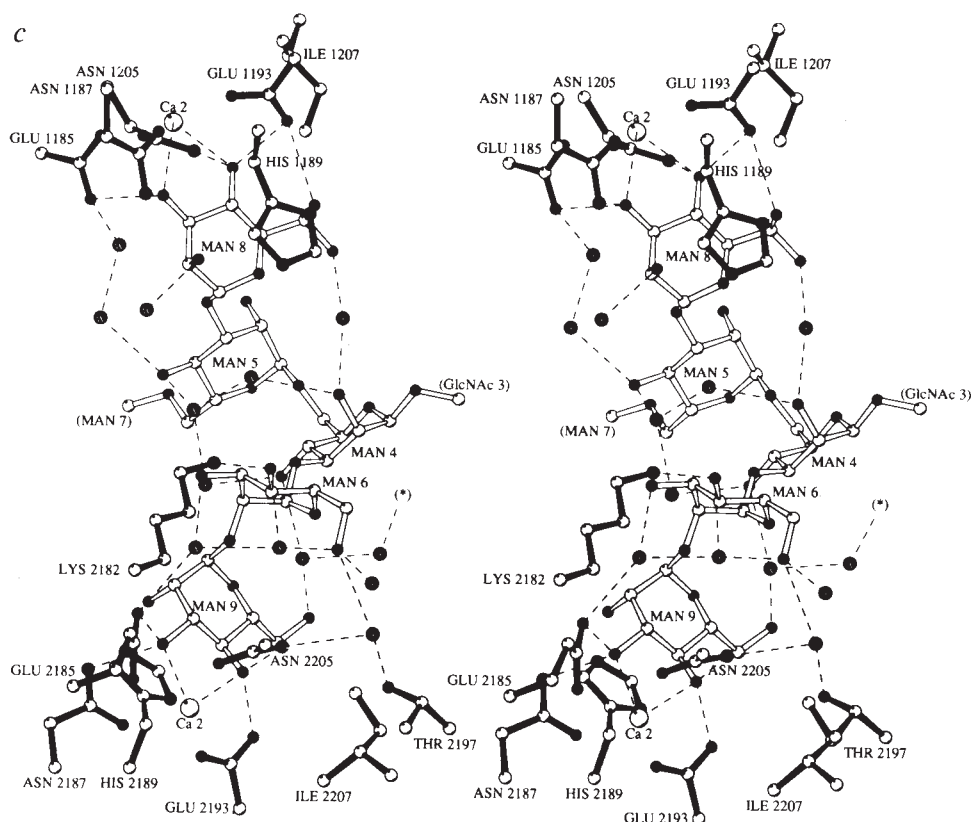
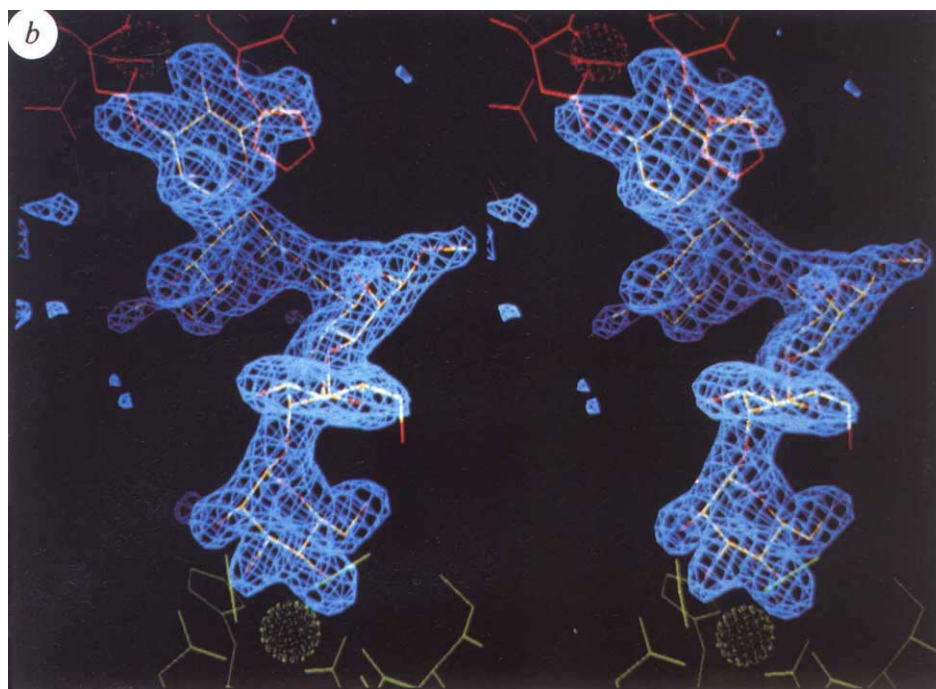
The inability of MBPs to bind galactose can be understood from the considerations outlined above. D-Galactose has an axial 4- and equatorial 2- and 3-OH groups. If the galactose ring is superimposed on the mannose ring at matching ring positions, its 4-OH group cannot form the coordination and hydrogen bonds made by the equatorial 4-OH of mannose, and is sterically excluded by C β of His 189 (Fig. 4*c*). Unlike L-fucose, the 2- and 3-OH groups of the D-galactose ring twisted with the opposite handedness with respect to the 3- and 4-OH groups of D-mannose, so they cannot form the same interactions with the protein. Interestingly, those C-type lectins that recognize mannose or glucose contain glutamate and asparagine at the positions equivalent to Glu 185 and Asn 187 of MBP-A, whereas those that recognize galactose have glutamine and aspartate at the corresponding positions, effectively reversing the position of the side-chain amide and carboxylate groups. The accompanying paper²⁹ tests the hypothesis that these residues are the primary determinants of mannose/glucose or galactose specificity in the C-type lectins.

Oligosaccharide structure and binding

Comparison of the glycosidic bond torsion angles observed in the Man₆GlcNAc₂Asn-CRD complex with those derived from

NMR measurements of the same asparaginyl-oligosaccharide (Table 2) reveals that the crystal lattice incorporates an oligosaccharide conformation from the ensemble that exists in solution, and does not impose a non-native structure upon the carbohydrate. Similar agreement between a crystallographically observed and NMR-derived oligosaccharide structure has been noted in crystals of *E. corallodendron* lectin, in which the N-linked oligosaccharide of the protein participates in crucial lattice interactions¹⁹. As NMR data do not uniquely determine the structure, the data are used as constraints on possible conformations obtained from energy calculations to arrive at a solution structure. All but one of the angles observed in the present structure fall in the same energy minima as those in solution^{30,31} (Table 2), and are also similar to those of Man₆GlcNAc₂ in solution³². The only significant discrepancy involves the ω angle of the internal $\alpha(1, 6)$ -linked Man 5, which is known to be a major site of conformational variability in N-linked oligosaccharides^{31,33}. The $\omega = +60^\circ$ rotamer appears to predominate in solution³¹, and the preference for this rotamer is hypothesized to be due to formation of a hydrogen bond between the 2-OH of Man 8 and the carbonyl oxygen of GlcNAc 3. The difference in energy between the $\omega = +60^\circ$ and $\omega = -60^\circ$ rotamers is small, however, and the interactions between Man 8 and the protein may favour incorporation of the minor conformer by compensating for the loss of any favourable interactions within the oligosaccharide. It should also be noted that the energetic contributions of bound water molecules, which are difficult to incorporate into the solution structure calculations, may affect the detailed conformation of the oligosaccharide.

As observed in other oligosaccharide structures^{16,34}, ordered water molecules stabilize the Man₆GlcNAc₂Asn oligosaccharide structure by linking together several mannose residues (Fig. 1*c*); no direct inter-mannose hydrogen bonds occur. In particular, the 6-OH of each of the two visible terminal mannose residues is linked by one water molecule to the central branching mannose (Man 4 in Fig. 1): Man 9 ($\alpha(1, 2)$ branch) to the 4-OH of Man 4, and the $\chi = 50^\circ$ conformer (Table 2) of Man 8 ($\alpha(1, 3)$ branch) to the 2-OH of Man 4. In the latter case, the hydrogen-bond



geometry is not especially favourable, which may explain why the 6-OH exists in two conformations, with the other conformer able to form hydrogen bonds with the protein (see above). These observations demonstrate that although the terminal mannoses bind similarly to the protein, the branch structure of the oligosaccharide, which puts the termini in different positions with respect to the other sugars in the chain, creates subtle differences in the mannose-CRD interactions.

Although the interactions of the two terminal mannoses with the CRD are virtually identical, the penultimate sugar of each branch forms different contacts with the CRD. With the possible exception of a poorly ordered lysine (182) side chain, water molecules mediate these interactions (Fig. 1c). For example, the 6-OH of the penultimate sugar of the $\alpha(1,2)$ branch (Man 6) is linked to Thr 197 O γ 1 through a water molecule, and another water molecule links the 3-OH to Glu 185 O ϵ 1. Similarly, a

chain of two water molecules links the 4-OH of the penultimate sugar of the $\alpha(1,3)$ branch (Man 5) to the protein. Even allowing for the requirements of crystal lattice packing, these observations suggest that the large number of sugar hydroxyl groups can be used in combination with ordered water molecules to enlarge the interaction surface of an oligosaccharide with a protein in a non-specific manner. Thus, although the only obligate features

of the CRD-oligosaccharide interaction appear to be those described for the terminal mannose residues, water-mediated interactions of the protein with the other sugars of the oligosaccharide appear to play a significant role in binding. The importance of ordered water molecules has also been noted in the binding of oligosaccharides to *L. ochrus* isolectin I (refs 34, 35).

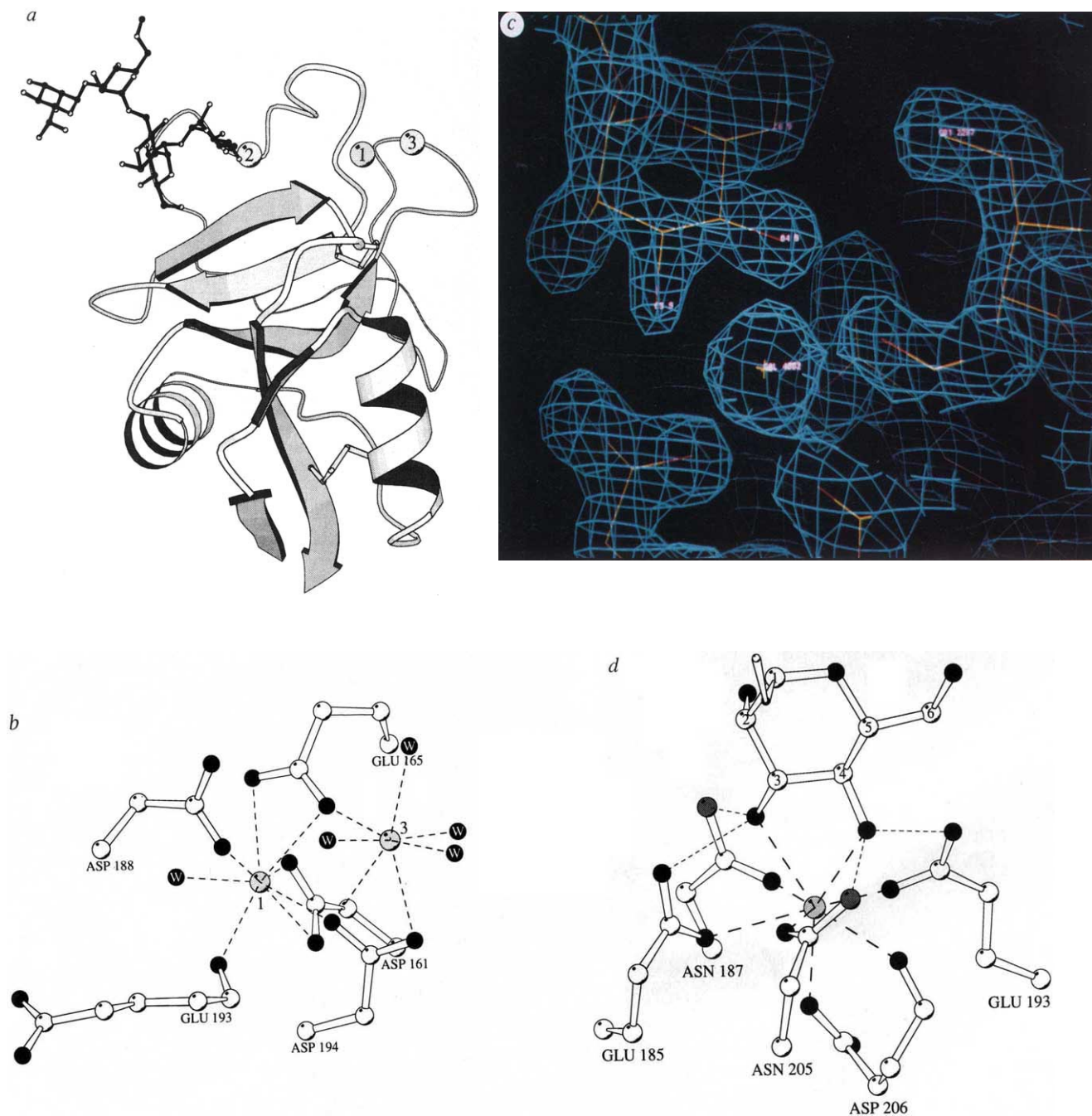


FIG. 2 Ca^{2+} and sugar binding to the CRD. *a*, View of the oligosaccharide binding to one CRD protomer. The CRD is shown in a ribbon representation; numbered spheres represent the three Ca^{2+} sites. *b*, Close-up of Ca^{2+} sites 1 and 3. White, black, and light grey spheres represent carbon, oxygen, and calcium, respectively. Black spheres marked 'W' are water molecules. Dashed lines denote calcium coordination bonds. *c*, Close-up of the $2F_o - F_c$ electron density map made from the final model (1.0σ contour) at calcium site 2, showing the terminus of the $\alpha(1,2)$ branch of the oligosaccharide. The orientation here and in *d* is similar to that shown in Fig. 1*c*. The packing of carbon 6 of mannose against Ile 207 C δ 1 can be seen near the top. *d*, Mannose binding at calcium site 2. The calcium is shown as a light grey

sphere. White, dark grey, and black spheres represent carbon, nitrogen, and oxygen, respectively. Thick dashed lines represent calcium coordination bonds; thin dashed lines represent hydrogen bonds. The α -glycosidic bond to the next sugar of the oligosaccharide (at carbon 1) has been cut off for clarity. Oxygen atoms from the side chains of Glu 185, Asn 187, Glu 193 and Asn 205, and the main-chain carbonyl oxygen of Asp 206, form the base of the pentagonal bipyramid coordination set of Ca^{2+} 2; a carboxylate oxygen of Asp 206 forms one apex of the bipyramid, and the 3- and 4-OH of mannose bisect the other apex. Panels *a*, *b*, and *d* were made with the program MOLSCRIPT⁴⁹.

C-type and other lectins discriminate among different carbohydrates with great specificity, despite their extremely weak affinities for monovalent carbohydrate ligands (K_d typically $\sim 10^{-5}$ to 10^{-3} mM). The strong and specific lectin-carbohydrate interactions that occur naturally at cell surfaces most probably arise through multivalent interactions, which magnify small differences in monovalent carbohydrate affinity into biologically significant effects. The precise, although limited, interactions seen between MBP and mannose help to explain how weak but specific binding occurs in the C-type lectins. The high-resolution picture of the binding of a single oligosaccharide to two different CRDs also suggests some of the structural principles likely to govern multivalent binding. In particular, the water-mediated, non-specific binding of MBP to internal sugars appears to complement binding to the terminal sugars of non-equivalent branches, and thus permits the specific recognition of ligands presented in different contexts on the cell surface.

It should be noted that the CRD-oligosaccharide complex described here cannot represent the spatial arrangement of CRDs with respect to cell-surface oligosaccharides *in vivo*. First, the glycopeptide used in this study appears to be a monovalent ligand for MBP in solution, as it inhibits binding of the bacterially expressed trimer to yeast with a K_i of 1.3 mM¹⁷. Also, protein-protein interactions in the lattice undoubtedly influence the respective positions of the carbohydrate-binding sites. Most importantly, as discussed above, the oligomeric structure of the MBP fragment studied here is an artefact of its preparation. The crystal structure of the trimeric fragment must be completed before the natural disposition of carbohydrate-binding sites and the possibilities for the multivalent interactions between MBP and oligomannose oligosaccharides can be understood. As multivalent binding often involves the clustering of lectin domains, the oligomeric structure of MBP and other C-type lectins remains a key unsolved problem in understanding the

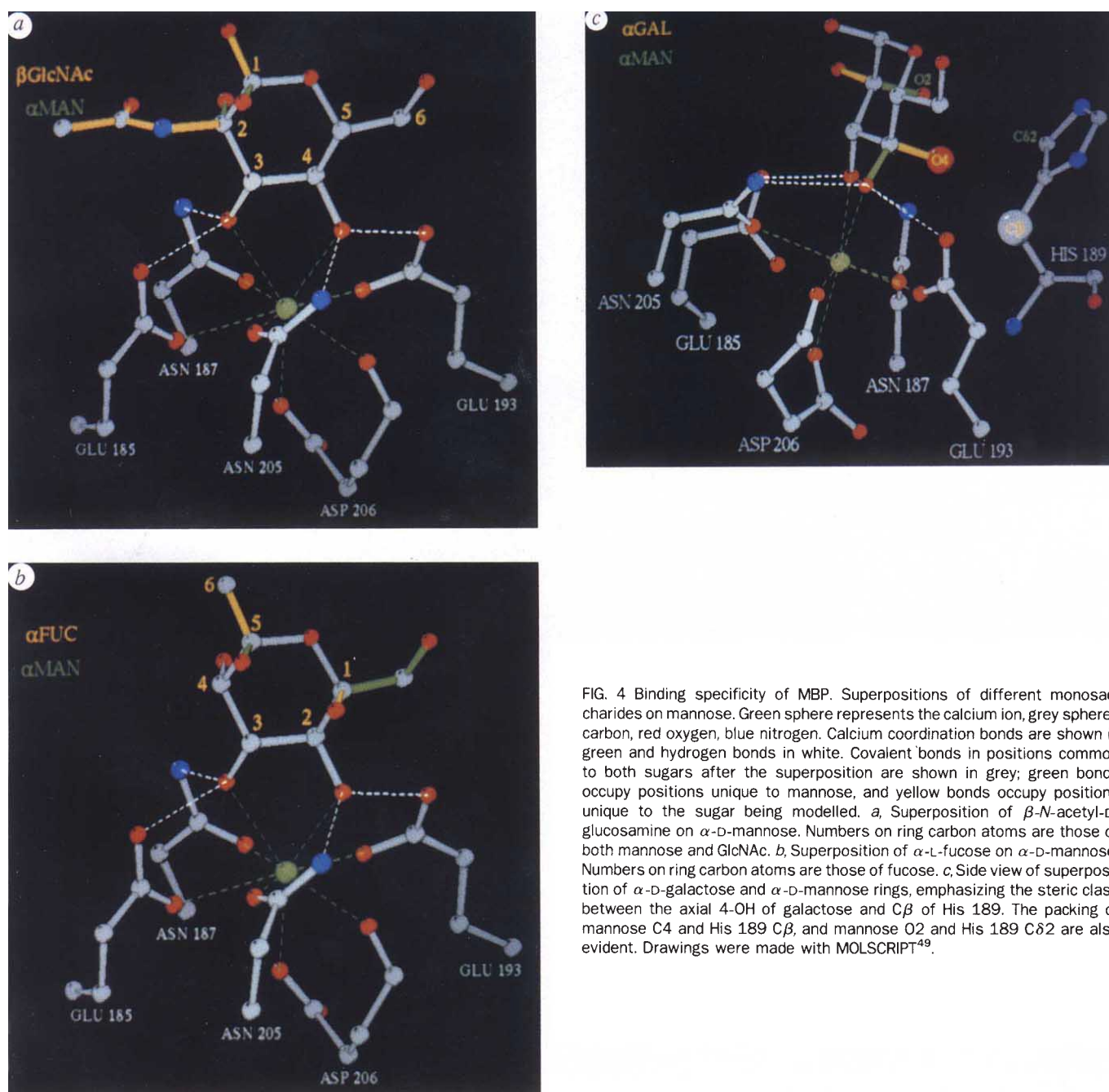


FIG. 4 Binding specificity of MBP. Superpositions of different monosaccharides on mannose. Green sphere represents the calcium ion, grey spheres carbon, red oxygen, blue nitrogen. Calcium coordination bonds are shown in green and hydrogen bonds in white. Covalent bonds in positions common to both sugars after the superposition are shown in grey; green bonds occupy positions unique to mannose, and yellow bonds occupy positions unique to the sugar being modelled. *a*, Superposition of β -N-acetyl-D-glucosamine on α -D-mannose. Numbers on ring carbon atoms are those of both mannose and GlcNAc. *b*, Superposition of α -L-fucose on α -D-mannose. Numbers on ring carbon atoms are those of fucose. *c*, Side view of superposition of α -D-galactose and α -D-mannose rings, emphasizing the steric clash between the axial 4-OH of galactose and C β of His 189. The packing of mannose C4 and His 189 C β , and mannose O2 and His 189 C δ 2 are also evident. Drawings were made with MOLSCRIPT⁴⁹.

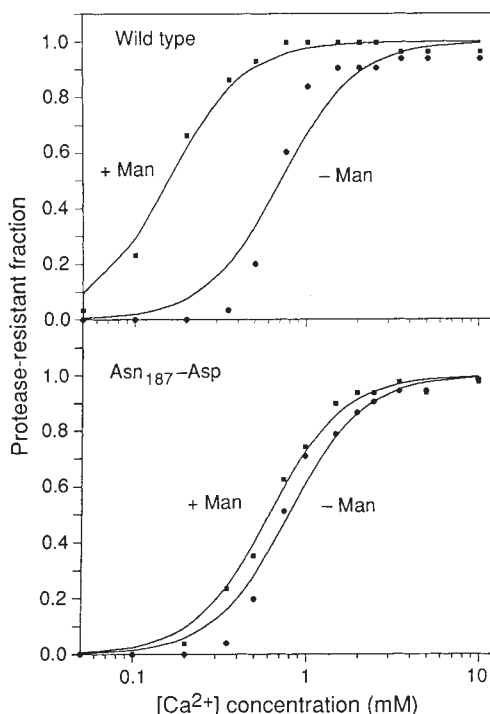


FIG. 3 Loss of mannose- but not Ca^{2+} -binding activity in the mutant Asn 187 $\rightarrow$ Asp. Ca^{2+} -dependent resistance to proteolysis, used to assess formation of the native, Ca^{2+} -ligated conformation, was determined by digestion of wild-type (upper) and mutant (lower) CRDs with subtilisin followed by gel electrophoresis. Second-order k_{Ca} values are 0.8 ± 0.1 and 0.85 ± 0.05 mM for wild-type and mutant, respectively, in the absence of mannose, and 0.20 ± 0.05 and 0.60 ± 0.05 mM for wild-type and mutant, respectively, in the presence of 200 mM mannose (average of three experiments).

METHODS. After renaturation of a bacterially expressed fragment containing the neck and CRD of MBP¹⁵, wild-type and mutant CRDs were prepared by proteolysis without affinity chromatography. Ca^{2+} concentration was adjusted to 50 mM, and digestion with $40 \mu\text{g ml}^{-1}$ of subtilisin was performed for 1 h at 37 °C. The reaction was stopped by addition of phenylmethylsulphonyl fluoride to a concentration of $300 \mu\text{g ml}^{-1}$, followed by dialysis against water, lyophilization and purification by reverse-phase chromatography on a C3 column¹⁵. Analytical digestion of the isolated CRDs was performed and quantified as described¹⁵.

mechanism by which these proteins mediate carbohydrate recognition *in vivo*. Nonetheless, the identical mode of binding by the MBP CRD to different terminal mannoses of Man_6Glc -

NAC_2Asn observed here demonstrates that multivalent binding can be achieved through specific recognition of sugars presented on non-equivalent branches of an oligosaccharide. □

Received 30 July; accepted 28 September 1992.

- Drickamer, K. *J. biol. Chem.* **263**, 9557–9560 (1988).
- Ashwell, G. & Harford, J. A. *Rev. Biochem.* **51**, 531–554 (1982).
- Spiess, M. *Biochemistry* **29**, 10009–10018 (1990).
- Taylor, M. E., Conary, J. T., Lennartz, M. R., Stahl, P. & Drickamer, K. *J. biol. Chem.* **265**, 12156–12162 (1990).
- Bevilacqua, M. P., Stengelin, S., Gimbrone, M. A. J. & Seed, B. *Science* **243**, 1160–1165 (1989).
- Johnston, G. I., Cook, R. G. & McEver, R. P. *Cell* **56**, 1033–1044 (1989).
- Lasky, L. A. *et al. Cell* **56**, 1045–1055 (1989).
- Siegelman, M. H., van de Rijn, M. & Weissman, I. L. *Science* **243**, 1165–1172 (1989).
- Halberg, D. F., Proulx, G., Doege, K., Yamada, Y. & Drickamer, K. *J. biol. Chem.* **263**, 9486–9490 (1988).
- White, R. T. *et al. Nature* **317**, 361–363 (1985).
- Ikedo, K., Sannoh, T., Kawasaki, N., Kawasaki, T. & Yamashina, I. *J. biol. Chem.* **262**, 7451–7454 (1987).
- Kuhlman, M., Joiner, K. & Ezekowitz, R. A. B. *J. exp. Med.* **169**, 1733–1745 (1989).
- Drickamer, K., Dordal, M. S. & Reynolds, L. *J. biol. Chem.* **261**, 6878–6886 (1986).
- Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. *Science* **254**, 1608–1615 (1991).
- Weis, W. I., Crichton, G. V., Murthy, H. M. K., Hendrickson, W. A. & Drickamer, K. *J. biol. Chem.* **266**, 20678–20686 (1991).
- Wright, C. S. *J. biol. Chem.* **267**, 14345–14352 (1992).
- Lee, R. T. *et al. J. biol. Chem.* **266**, 4810–4815 (1991).
- Derewenda, Z. *et al. EMBO J.* **8**, 2189–2193 (1989).
- Shaanan, B., Lis, H. & Sharon, N. *Science* **254**, 862–866 (1991).
- Bourne, Y. *et al. Proteins* **8**, 365–376 (1990).
- Quioco, F. A. *Rev. Biochem.* **55**, 287–315 (1986).
- Vyas, N. K. *Curr. Opin. struct. Biol.* **1**, 732–740 (1991).
- Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
- Quesenberry, M. S. & Drickamer, K. *J. biol. Chem.* **267**, 10831–10841 (1992).
- Lee, Y. C. & Lee, R. T. *In The Glycoconjugates Vol. 4* (ed. Horowitz, M. I.) 57–84 (Academic, Orlando, 1982).

- Brandley, B. K., Swiedler, S. J. & Robbins, P. W. *Cell* **63**, 861–863 (1990).
- Siegelman, M. *Curr. Biol.* **1**, 125–128 (1991).
- Springer, T. A. & Lasky, L. A. *Nature* **349**, 196–197 (1991).
- Drickamer, K. *Nature* **360**, 183–186 (1992).
- Brisson, J.-R. & Carver, J. P. *Biochemistry* **22**, 3671–3680 (1983).
- Brisson, J.-R. & Carver, J. P. *Biochemistry* **22**, 3680–3686 (1983).
- Homans, S. W., Pastore, A., Dwek, R. A. & Rademacher, T. W. *Biochemistry* **26**, 6649–6655 (1987).
- Homans, S. W. *et al. Biochemistry* **25**, 6342–6350 (1986).
- Bourne, Y., Rougé, P. & Cambillau, C. *J. biol. Chem.* **267**, 197–203 (1992).
- Bourne, Y., Rougé, P. & Cambillau, C. *J. biol. Chem.* **265**, 18161–18165 (1990).
- Brünger, A. T. *X-PLOR Manual, Version 3.0* (Yale Univ., New Haven, 1992).
- Brünger, A. T. *Acta crystallogr.* **A46**, 46–57 (1990).
- Bilderback, D. *et al. Nucl. Instrum. Meth.* **A266**, 636–644 (1988).
- Kabsch, W. *J. appl. Crystallogr.* **21**, 67–71 (1988).
- CCP4. *The SERC (UK) Collaborative Computing Project No. 4, A Suite of Programs for Protein Crystallography* (Daresbury Laboratory, Warrington WA4 4AD, UK, 1979).
- Sheriff, S. & Hendrickson, W. A. *Acta crystallogr.* **A43**, 118–121 (1987).
- Engh, R. A. & Huber, R. *Acta crystallogr.* **A47**, 392–400 (1991).
- Brünger, A. T. *Nature* **355**, 472–475 (1992).
- Weis, W. I., Brünger, A. T., Skehel, J. J. & Wiley, D. C. *J. molec. Biol.* **212**, 737–761 (1990).
- Peters, D. & Peters, J. *J. molec. Struct.* **85**, 107–123 (1981).
- Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535–542 (1977).
- Imberty, A., Gerber, S., Tran, V. & Pérez, S. *Glycoconjugate J.* **7**, 27–54 (1990).
- Brisson, J.-R. & Carver, J. P. *Biochemistry* **22**, 1362–1368 (1983).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).

ACKNOWLEDGEMENTS. We thank CHESS for synchrotron beam time, N. Ruiz for assistance in protein purification, P. Kwong for help with data collection and A. Brünger for discussions. W.I.W. was a Life Sciences Research Foundation Fellow of the Howard Hughes Medical Institute. Supported by an American Cancer Society Faculty Salary award (K.D.) and by grants from the National Institutes of Health (K.D.) and the National Science Foundation (W.A.H.).

X-ray flares from runaway pair production in active galactic nuclei

J. G. Kirk & A. Mastichiadis*

Max-Planck-Institut für Kernphysik, 6900 Heidelberg 1, Germany

ACTIVE galactic nuclei (AGNs) exhibit high luminosity with rapid variability, especially in X-ray emission, for which the luminosity can be greater than that of a normal galaxy, and the variability timescale implies an emitting region in some cases smaller than one light hour¹. The hard X-ray spectrum of AGNs is nonthermal, probably arising from an electron–positron pair cascade, with some emission reflected off relatively cold matter^{2,3}. Energy can be pumped into a cascade by any process that produces relativistic pairs, but because electrons and positrons are hard to accelerate efficiently, there has been interest in models in which protons are accelerated^{4,5}, and create relativistic electrons on interaction with a local radiation field^{6–8}. Here we show that a sufficient column density of protons can lead to runaway pair production: photons generated by the relativistic pairs are the targets for the protons to produce more pairs. This process can produce X-ray flares with the observed characteristics, and our model predicts the maximum ratio of luminosity to source size ('compactness') as well as their spectrum in the early phases. The same mechanism may also be able to create the knots of synchrotron-radiating pair plasma seen in sources such as 3C273.

There are several feedback mechanisms that could operate in the particle acceleration zone of an AGN^{9–11}. The one that we propose exists when sufficiently relativistic protons are present for the following chain reaction to occur: a single photon collides with a relativistic proton producing an electron–positron pair, which then radiates many photons by the synchrotron process. Most of these escape, but when the optical thickness of energetic protons within the source exceeds a certain critical value, more than one of the emitted photons may be intercepted to produce another pair, leading to a runaway. The number of photons then increases rapidly, and the X-ray spectrum from the linear phase of the instability has an index close to those observed in most AGNs, given the $n_p \propto \gamma^{-2}$ distribution expected of protons accelerated at a strong shock front (here n_p is the number density of protons and γ is the Lorentz factor). In the nonlinear phase, when the photon density has built up, a combination of proton–photon pair production and photon-induced pion production may cause effective proton cooling. We propose that the X-ray flaring observed in some AGNs could be a result of this chain reaction, provided that the quiescent or pre-flare phase does not support an electromagnetic cascade. The maximum energy that can be released in a flare is determined by the instability criterion, leading to a maximum theoretical value of the 'compactness' parameter¹² which is a function of the proton spectrum. In the case of 3C273, flaring is associated with the birth of relativistically moving knots in the jet¹³. We speculate that our mechanism, which predicts an interval between flares of a few years, may be responsible for the creation of the pair plasma subsequently seen as a knot.

We will consider a distribution of relativistic protons confined in a volume of characteristic radius R in which a magnetic field of strength B is present. To investigate the stability of this system we suppose that both photons and electrons are initially present only as small perturbations and that the density of the background plasma is small enough for proton–proton reactions to be unimportant. The processes remaining after linearization of the kinetic equations describing the electron and photon number densities are synchrotron emission, Bethe–Heitler pair produc-

tion by a photon in the field of a relativistic proton, and photo-pion production leading to pairs, photons, neutrinos and neutrons. The processes usually considered in electromagnetic cascades, inverse Compton scattering and photon–photon pair production, are quadratic in the perturbations and will be neglected. There remain two feedback loops which could lead to instability: the pair-production/synchrotron, and the π^+ production/synchrotron loops. For the pair-production loop, the kinetic equation governing the number density $n_e(\gamma, t)$ of electrons and positrons as a function of Lorentz factor γ reads:

$$\frac{\partial n_e(\gamma, t)}{\partial t} - \frac{4}{3} l_B \frac{\partial}{\partial \gamma} [\gamma^2 n_e(\gamma, t)] - 2n_p(\gamma) \int_0^\infty dx n_\gamma(x, t) \sigma_{pe}(x\gamma) = 0 \quad (1)$$

Here $\sigma_{pe}(x)$ is the pair-production cross-section in units of the Thomson cross-section σ_T as a function of the photon energy $xm_e c^2$ in the proton's rest frame; $n_e(\gamma, t) d\gamma$, $n_p(\gamma) d\gamma$ and $n_\gamma(x, t) dx$ are the differential numbers of electrons, protons and photons respectively in the volume element $\sigma_T R$; time t is measured in units of the light crossing time of the system R/c ; and we have introduced the 'magnetic compactness'

$$l_B = \left(\frac{U_B}{m_e c^2} \right) \sigma_T R \quad (2)$$

in which $U_B = B^2/8\pi$ is the magnetic energy density and m_e the electron mass. The second term in equation (1) describes synchrotron cooling and the third is a source term from proton–photon pair production, with the implied assumption that pair production by a proton of energy $\gamma m_p c^2$ results in a positron and electron each with energy $\gamma m_e c^2$. This is accurate because most interactions occur near threshold, as we show below. To accompany equation (1) there is a kinetic equation for the photon distribution which includes photon escape through the boundaries of the emission region and synchrotron emission (the absorption of photons in pair production is negligible):

$$\frac{\partial n_\gamma(x, t)}{\partial t} + n_\gamma(x, t) - \frac{2}{3} l_B b^{-3/2} x^{-1/2} n_e(\sqrt{x/b}, t) = 0 \quad (3)$$

where $b = B/B_c$ denotes the magnetic field in units of the critical field $B_c = m_e^2 c^3/(e\hbar) = 4.414 \times 10^{13}$ G and we have approximated the synchrotron radiation of a single electron of energy $\gamma m_e c^2$ as monochromatic at a (dimensionless) frequency $x = b\gamma^2$. This simple approximation allows a Laplace transformation of equations (1) and (3), leading to the result that both the electron and photon densities behave at large t as $\exp(st)$, where s is the largest solution of the eigenvalue problem:

$$(s+1)n_\gamma(x, s) = b^{-1/2} x^{-3/2} \exp\left(\frac{-3s}{4l_B} \sqrt{\frac{b}{x}}\right) \times \int_0^\infty dx' n_\gamma(x', s) \int_{\sqrt{x/b}}^\infty d\gamma n_p(\gamma) \times \sigma_{pe}(x'\gamma) \exp\left(\frac{3s}{4l_B \gamma}\right) \quad (4)$$

(The Laplace-transformed densities are indicated by simply changing the argument from t to s .) If there are no protons above $\gamma = \gamma_{\max}$, then for $x > b\gamma_{\max}^2$ the photons escape freely: $s = -1$. Because only growing solutions are of interest, we must have $x' < b\gamma_{\max}^2$ in the first integral, and, because of the threshold condition for pair production, $x'\gamma > 2$, we find a critical value of $\gamma_{\max}$ given by

$$\gamma_{\text{crit}} = (2/b)^{1/3} \quad (5)$$

To have feedback, protons of energy larger than $\gamma_{\text{crit}} m_p c^2$ must exist. This condition is actually quite weak. For a field of 10^3 G, we require protons of energy only a few teraelectronvolts (TeV);

* Present address: NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

the corresponding synchrotron photons are in the X-ray range. To produce an instability, however, the feedback must be strong enough to overcome photon escape. Taking a proton distribution $n_p(\gamma) = n_0/\gamma^\beta$, it is easy to find the condition for marginal stability by setting $s=0$ in the integral eigenvalue problem (equation (4)). We find runaway pair production when

$$n_0 > n_{\text{crit}} \equiv \left(\frac{2}{3}\beta - 1\right) b^{1-\beta/3} \left[\int_0^\infty dy \sigma_{\text{pe}}(y) y^{-1-\beta/3} \right]^{-1} \quad (6)$$

and the corresponding photon distribution, which is just the eigenfunction of eigenvalue $s=0$, is

$$n_\gamma(x, s=0) \propto x^{-1-\beta/3} \quad (7)$$

The integral in equation (6) is dominated by values of y close to threshold, provided the proton spectrum is not too flat, and has the value 0.002 (0.0035) for $\beta=2$ ($\beta=1.65$). The associated photon spectrum is $\propto x^{-1.67}$ ($\propto x^{-1.55}$), and is rather insensitive to the precise value of β .

A result analogous to equation (6) can be found for the photo-pion process if the $p\gamma$ cross-section is approximated by a delta function $\sigma_{p\gamma}(y) \approx \sigma_0 \sigma_T \delta(y - y_0)$ (where $y = \gamma x$ is the dimensionless photon frequency in the proton rest frame and γ the proton's Lorentz factor). If the interaction results in the injection of a total of M_e electrons and positrons each of energy $a\gamma m_e c^2$, a straightforward calculation leads to

$$n_{\text{crit}}^{(\pi)} = \frac{2a^{1-2(\beta/3)} y_0^{1+(\beta/3)} b^{1-(\beta/3)}}{3M_e \sigma_0} (2\beta - 3) \quad (8)$$

For π^+ production, we have $y_0 \approx 600$, $\sigma_0 \approx 0.25$, $M_e = 1$ and $a \approx 600/M_e$, so that the threshold for instability is roughly two orders of magnitude higher than for the pair production/synchrotron loop.

A realistic proton spectrum cannot be a simple power law at all energies as assumed in equation (6): at low energies, complications arise with transrelativistic particles, and at high energies cooling must become important at some stage. This leads us to

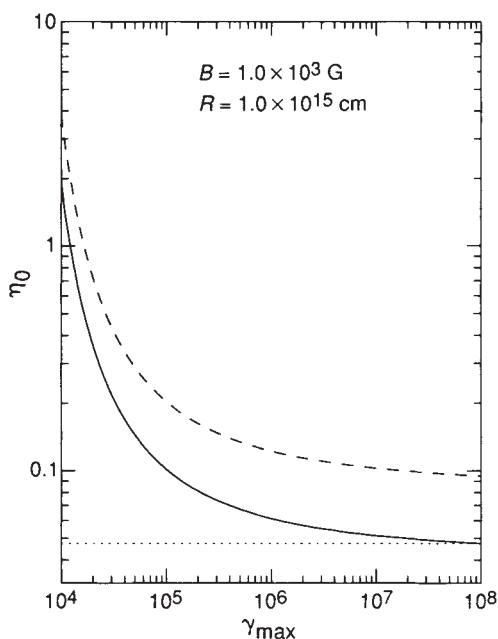


FIG. 1 The value of n_0 for marginal stability (solid line) of a truncated power law distribution (equation (9)) of index $\beta=2$ as a function of γ_{max} , with $\gamma_{\text{min}}=1$ and a magnetic field of 10^3 G. The dotted line shows the asymptotic limit of the marginal stability condition for large γ_{max} . The dashed line depicts the value at which the growth time of the instability equals the light crossing time of the source, taken to be of size 10^{15} cm.

approximate the distribution by a truncated power-law:

$$n_p = n_0/\gamma^\beta \quad \text{for } \gamma_{\text{min}} < \gamma < \gamma_{\text{max}} \\ = 0 \quad \text{otherwise} \quad (9)$$

Figure 1 shows the results of a numerical solution of equation (4) using this distribution with $\beta=2$. For n_0 above the marginal instability line, photons and pairs grow exponentially with an e-folding time which is shorter for larger n_0 , reaching the light crossing time ($s=1$) at the dashed line. The parameters chosen for Fig. 1 are such that the light crossing time is a fraction of a day, roughly the size expected of the X-ray emitting region of a not too rapidly variable AGN (such as 3C273) and the magnetic field is taken to be 10^3 G (ref. 14). The magnetic compactness is then $l_B = 32$. The condition of marginal instability is, however, only a weak function of the magnetic field, as can be seen from equation (6), and is independent of the size of the emitting region; it is the growth rate of the instability and the position of the dashed line that are functions of l_B and, through this quantity, depend on both B and R .

Variability of X-ray flux is common among AGNs (well known examples include 3C273 (ref. 15), NGC4151 (refs 16, 17), NGC4051 (ref. 18) and NGC5548 (ref. 19)). Some sources exhibit flaring, for which the process described above can act as a triggering mechanism. Assume that the source injects energetic protons continuously. As long as the column density of protons stays below the critical value given (roughly) by equation (6), the density of photons cannot, in the absence of a cascade, maintain itself and energy is extracted from the protons only if there is an external photon source, such as a hot accretion disk. But once the critical column density is exceeded, runaway production of photons and pairs occurs. The subsequent behaviour is more complicated. Photon-photon pair production and inverse Compton scattering may become important, setting off an electromagnetic cascade, and proton cooling by photo-pion production may occur. Nevertheless, the energy content in protons E_{max} just before such a flare is well defined and, because this energy cannot be liberated on a timescale faster than R/c , we can estimate the maximum compactness. Taking $\beta=2$, we find

$$l_{\text{max}} = \frac{(E_{\text{max}} c/R) \sigma_T}{4\pi R m_p c^3} \approx n_0 \left[\frac{m_p \log(\gamma_{\text{max}}/\gamma_{\text{min}})}{3m_e} \right] \quad (10)$$

According to Fig. 1, $l_{\text{max}} \approx 465(B/10^3 \text{ G})^{1/3}$ if we take $\gamma_{\text{max}}/\gamma_{\text{min}} \sim 10^7$.

In the special case of 3C273, extensive multi-frequency observations are available. Apparently superluminal motion of radio emitting knots of plasma is seen emerging from the central source and moving out along the direction of the jet²⁰⁻²². In two cases a knot seems to have been born at about the same time as an X-ray flare of slightly steeper spectrum than average¹³. The spectrum of this source is rather flat ($\alpha \approx 0.5$), close to that emitted by monoenergetically injected electrons cooling by synchrotron or inverse Compton radiation by scattering off soft photons. The steepening that would be produced by an electromagnetic cascade²³ seems to be absent. In the present picture in which proton acceleration is responsible for the electron injection, a flat spectrum can arise when the acceleration mechanism is saturated by interaction with the strong flux of ultraviolet photons which are presumably emitted by the accretion disk^{8,24}. Accelerated protons tend to accumulate at the energy at which the ultraviolet losses set in, and, as well as producing many electrons and positrons of this characteristic energy, are themselves injected into the source essentially monoenergetically²⁵. The resulting spectrum of cooling electrons is $n_e \propto \gamma^{-2}$, whereas the protons cool primarily by interacting with photons of energy higher than those of the ultraviolet bump. Interactions with the background plasma can also lead to proton cooling, but the importance of this process depends on whether acceleration occurs in a low-density region (such as a corona

above the accretion disk). If we assume protons are neither accreted into the black hole nor ejected, then they simply accumulate in the source with a distribution in energy which can be approximated by equation (9) with $\gamma_{\max}$ given by the Lorentz factor at injection ($\sim 10^7$) and $\gamma_{\min}(t)$ the lowest value to which they have cooled. The lowest value of γ at which protons can still contribute to the pair production instability is $\gamma_{\min} = 2/\sqrt{b\gamma_{\max}} \approx 130$, and we estimate that this value is reached by a proton after cooling for a few years. In our model, the most energetic explosions should occur after such an accumulation time. The proton spectral index appropriate to the distribution of equation (9) is that which gives the observed X-ray spectrum of the flare ($\alpha \approx 0.55$) in the linear phase of the instability, $\beta = 1.65$. For a spherical source of size R_{15} (in units of 10^{15} cm) and a magnetic field of B_3 (in units of 10^3 G) we can use equation (6) to estimate the total energy involved

$$E_{\max} = 3.5 \times 10^{51} R_{15}^2 B_3^{0.45} (\gamma_{\max}/10^7)^{0.35} \text{ erg} \quad (11)$$

This represents the maximum possible energy release in a flare according to our model and is consistent with the luminosity of the X-ray flare observed by Ginga¹⁵ in December 1987, if we assume that it extended over a few days.

This picture leaves open the question of the role played by photo-pion creation, which is responsible for the widely predicted emission of relativistic neutrons and high-energy neutrinos^{8,26–28}. It will also be important to understand how the instability evolves and saturates in the nonlinear phase. Preliminary numerical calculations taking nonlinear effects into account^{10,11,29} depend sensitively on the details of the particle acceleration model adopted, about which there is considerable uncertainty. Nevertheless, the analytical results concerning spectral index and maximum compactness are compatible with the

observed X-ray spectra of many variable AGNs, and the instability may influence the birth of the superluminal knots in 3C273. $\square$

Received 30 July; accepted 4 September 1992.

1. Done, C. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **240**, 81–102 (1989).
2. Pounds, K. A., Nandra, K., Stewart, G. C., George, I. M. & Fabian, A. C. *Nature* **344**, 132–133 (1989).
3. Zdziarski, A. A., Ghisellini, G., Svensson, R., Fabian, A. C. & Done, C. *Astrophys. J.* **363**, L1–L4 (1990).
4. Protheroe, R. J. & Kazanas, D. *Astrophys. J.* **265**, 620–624 (1983).
5. Kazanas, D. & Ellison, D. C. *Astrophys. J.* **304**, 178–187 (1986).
6. Sikora, M., Kirk, J. G., Begelman, M. C. & Schneider, P. *Astrophys. J.* **320**, L81–L85 (1987).
7. Begelman, M. C., Rudak, B. & Sikora, M. *Astrophys. J.* **362**, 38–51 (1990).
8. Szabo, A. P. & Protheroe, R. J. in *Particle Acceleration in Cosmic Plasma* (eds Gaisser, T. & Zank, G. P.) (American Institute of Physics, in the press).
9. Bisnovaty-Kogan, G. S., Zel'dovich, Ya. B. & Sunyaev, R. A. *Sov. Astr.* **15**, 17–22 (1971).
10. Stern, B. & Svensson, R. in *Relativistic Hadrons in Cosmic Compact Objects* (eds Zdziarski, A. A. & Sikora, M.) (Springer, Berlin, 1991).
11. Stern, B. E., Sikora, M. & Svensson, R. in *Testing the AGN paradigm* (eds Holt, S. S., Neff, S. G. & Megan Urry, C.) 313–316 (American Institute of Physics, New York, 1992).
12. Guilbert, P. W., Fabian, A. W. & Rees, M. *Mon. Not. R. astr. Soc.* **205**, 593–603 (1983).
13. Walter, R. & Courvoisier, T. J.-L. *Astr. Astrophys.* **258**, 255–264.
14. Camenzind, M. in *Variability of Active Galaxies* (eds Duschl, W. J., Wagner, S. J. & Camenzind, M.) 201–210 (Springer, Berlin, 1991).
15. Turner, M. J. L. *et al.* *Mon. Not. R. astr. Soc.* **244**, 310–321 (1990).
16. Perola, G. C. *et al.* *Astrophys. J.* **306**, 508–521 (1986).
17. Yaqoob, T. & Warwick, R. S. *Mon. Not. R. astr. Soc.* **248**, 773–786 (1991).
18. Lawrence, A., Watson, M. G., Pounds, K. A. & Elvis, M. *Nature* **325**, 694–696 (1987).
19. Walter, R. & Courvoisier, T. J.-L. *Astr. Astrophys.* **233**, 40–52 (1990).
20. Zenzus, J. A., Unwin, S. C., Cohen, M. H. & Biretta, J. A. *Astr. J.* **100**, 1777–1784 (1990).
21. Krichbaum, T. P. *et al.* *Astr. Astrophys.* **237**, 3–11 (1990).
22. Bååth, L. B. *et al.* *Astr. Astrophys.* **241**, L1–L4 (1991).
23. Svensson, R. *Mon. Not. R. astr. Soc.* **227**, 403–451 (1987).
24. Sikora, M. & Begelman, M. C. in *High Energy Neutrino Astrophysics* (ed. Stenger, V. J.) (World Scientific, in the press).
25. Schlickeiser, R. *Astr. Astrophys.* **136**, 227–236 (1984).
26. Kirk, J. G. & Mastichiadis, A. *Astr. Astrophys.* **213**, 75–79 (1989).
27. Mannheim, K. & Biermann, P. L. *Astr. Astrophys.* **221**, 211–220 (1989).
28. Stecker, F. W., Done, C., Salamon, M. H. & Sommers, P. *Phys. Rev. Lett.* **66**, 2697–2700 (1991).
29. Mastichiadis, A. & Kirk, J. G. in *High Energy Neutrino Astrophysics* (ed. Stenger, V. J.) (World Scientific, in the press).

ACKNOWLEDGEMENTS. We thank A. Atoyan, M. Begelman, T. Courvoisier and M. Sikora for discussion.

Electron density fluctuations in the local interstellar bubble

J. A. Phillips* & Andrew W. Clegg†

* Owens Valley Radio Observatory, California Institute of Technology, 105-24, Pasadena, California 91125, USA

† Naval Research Laboratory, Code 4213, Washington DC 20375, USA

THE interstellar medium in our vicinity forms an elongated cavity of X-ray-emitting gas, extending up to 200 pc from the Sun^{1,2}. The gas inside this local bubble is very hot ($\sim 10^6$ K) and tenuous (electron density $n_e \approx 0.005 \text{ cm}^{-3}$), and is typical of material in the coronal phase of the interstellar medium, thought to occur when hot stars or supernovae blow out holes in cooler and denser interstellar gas. Little is known about turbulent density fluctuations in the coronal gas, but its physical state is important in understanding cosmic ray confinement and the energy balance of interstellar material, as well as the scattering of radio emission from compact sources^{3,4}. To probe turbulence in the local interstellar medium, we have made scintillation measurements at 50 MHz (6 m wavelength) of emission from the nearby pulsar 0950+08, which happens to lie near the edge of the local bubble. From the scintillation bandwidth we deduce the amplitude of the electron-density fluctuation spectrum in the coronal gas alone, and find it to be an order of magnitude lower than for any previously measured interstellar line of sight. We conclude that the interior of the bubble is relatively quiescent and that high levels of plasma turbulence—if they exist—must be localized to the cavity boundary.

Along many lines of sight through the interstellar medium (ISM), electron-density turbulence is well described by a power-law model of spatial density fluctuations⁵

$$P(q) = C_n^2 q^{-\alpha} \quad 2\pi/l_o \leq q \leq 2\pi/l_i \quad (1)$$

where P is the power spectrum of spatial fluctuations, q is the wavenumber, C_n^2 is a normalizing constant, and l_i and l_o are the inner and outer cut-offs, respectively. Measurements of angular broadening of compact radio sources⁵ and fluctuations in pulsar dispersion measure⁶ give estimates for α close to the 'Kolmogorov' value, 11/3. α is nearly constant for most lines of sight through the ISM, but C_n^2 is a strong function of the viewing direction⁷. Values inferred from previous scintillation measurements^{8,9} range from $\sim 10^{-3.5} \text{ m}^{20/3}$ for nearby and high-latitude pulsars to $10^5 \text{ m}^{-20/3}$ in the H II complex NGC6334.

It has proved difficult to measure C_n^2 within the local bubble because there are few nearby pulsars with scintillation properties dominated by the X-ray emitting gas. One of the most promising pulsars for such studies is PSR0950+08 ($l = 228.9^\circ$, $b = +43.7^\circ$). The distance to PSR0950+08 is 127 ± 13 pc (ref. 10) and the X-ray cavity extends for at least 80 pc in the direction of the pulsar^{11,12}. Recent H I absorption measurements yield an upper limit of $\leq 10^{18} \text{ cm}^{-2}$ on the neutral hydrogen column density along the line of sight to PSR0950+08 (D. Frail, personal communication). The low hydrogen column density confirms that PSR0950+08 lies within the well known H I void^{13,14}. The void extends from the location of the Sun (near the edge of the bubble) out to ~ 200 pc in the direction of the third quadrant of the galaxy (see Fig. 4 of ref. 14).

We obtained scintillation data for PSR0950+08 near 50 and 130 MHz with the 305-m Arecibo telescope. Autocorrelation function (ACFs) of the received electric fields were formed by a 40-MHz correlation spectrometer. The correlator was gated synchronously with the pulsar period, allowing the accumulation of ACFs as a function of pulse phase. The ACFs were averaged over integration times of ~ 12 s and written to magnetic tape. They were then Fourier-transformed to yield power spectra as a function of pulse phase. The spectra were integrated over phase bins when the pulsar was 'on' and normalized to spectra accumulated in the 'off-pulse' phase bins.

The normalized 'on-pulse' spectra were summed over both

TABLE 1 Dynamic spectra and scattering parameters

Date	ν (MHz)	δt (s)	$\delta \nu$ (kHz)	T (min)	B (MHz)	$\Delta \nu_d$ (kHz)	τ_d (s)	$\log C_n^2$ ($\text{m}^{-20/3}$)
26 November 89	47	13	9.8	33	1.250	28 ± 2	279 ± 17	-4.6
01 January 90	51	13	4.9	14	0.625	49 ± 7	227 ± 32	-4.7
01 January 90	51	13	4.9	20	0.625	31 ± 4	372 ± 48	-4.5
27 January 91	51	12	4.9	20	1.250	19 ± 1	243 ± 15	-4.4

Symbol definitions: δt is the time resolution of the dynamic spectrum; $\delta \nu$ is the frequency resolution; T is the total observing time; B is the observing bandwidth; $\Delta \nu_d$ and τ_d are the diffractive scintillation bandwidth and timescales, respectively. The latter are defined as the half-width at half maximum of the dynamic spectrum's ACF, along the frequency axis ($\Delta \nu_d$) or the time axis (τ_d). The fractional errors in $\Delta \nu_d$ and τ_d were estimated from the formula⁸ $\epsilon = (\Delta \nu_d \tau_d / BT)^{1/2}$.

polarizations and used to construct the dynamic spectrum, which is the flux density of the pulsar as a function of time and frequency. Our dynamic spectra were typically composed of 100–200 time samples and 512 spectral channels. The total observing bandwidth varied between 0.625 and 5 MHz, depending on the date of observation. Table 1 lists the observing parameters for our 50-MHz observations and Fig. 1 shows a typical dynamic spectrum.

The 50-MHz dynamic spectra closely resemble the well known 'patchy' structure produced by diffractive interstellar scintillations (compare our Fig. 1 with Figs 4–6 in ref. 8). Following standard practice, we estimated the diffractive decorrelation timescale (τ_d) and frequency scale ($\Delta \nu_d$) from the two-dimensional ACFs of the dynamic spectra. The values of τ_d and $\Delta \nu_d$ listed in Table 1 correspond to the half-width of the ACF along the time axis and frequency axis, respectively.

The scintillation bandwidth of PSR0950+08 has been measured before, but at higher frequencies between 0.3 and 1.0 GHz (refs 8, 15). Extrapolating our 50-MHz measurements to 320 MHz ($\Delta \nu_d \propto \nu^{4.4}$), we find $\Delta \nu_d \approx 100$ MHz. For comparison, Cordes *et al.*⁸ reported $\Delta \nu_d \approx 1$ MHz at 320 MHz, a value much smaller than expected from the 50-MHz observations. The high-frequency measurements were almost certainly biased by radio-meter bandwidths much less than the scintillation bandwidth. Figure 5 of ref. 15, for example, shows under-resolved frequency structure in a 340-MHz dynamic spectrum of the pulsar. We observed PSR0950+08 at 130 MHz using radiometer bandwidths of 0.625 MHz, 1.25 MHz, 2.5 MHz and 5.0 MHz (the expected decorrelation bandwidth was ~ 2 MHz), and found that dynamic spectra with the smallest bandwidths consistently yielded underestimates of $\Delta \nu_d$. In contrast, there are many well resolved intensity maxima in the 50-MHz dynamic spectra, and we believe that these data provide the most accurate estimate of $\Delta \nu_d$ so far.

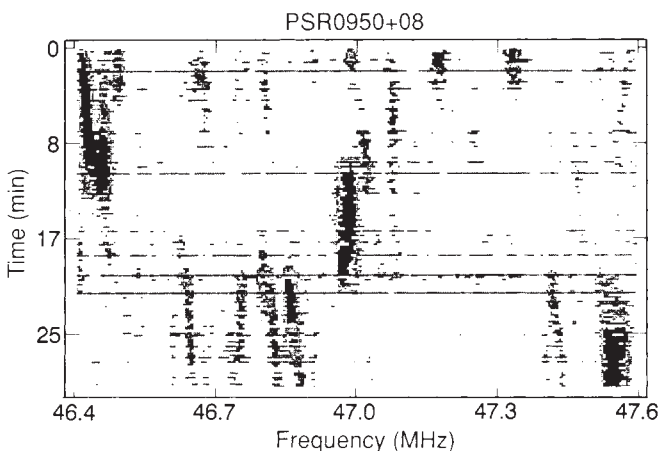


FIG. 1 Dynamic spectrum of PSR0950+08 obtained at 47 MHz on 1989 November 26. The grey scale represents flux density as a function of time and frequency, with black indicating strong flux density. The black horizontal lines are terrestrial interference.

Cordes *et al.*⁸ have discussed in detail the estimation of C_n^2 from decorrelation bandwidth measurements. On the basis of their work we adopted the estimator

$$\hat{C}_n^2 = 2 \times 10^{-6} \nu^{11/3} D^{-11/6} \Delta \nu_d^{-5/6} \text{ m}^{-20/3} \quad (2)$$

with ν in megahertz, D in parsec, and $\Delta \nu_d$ in kilohertz. We estimate the errors in $\hat{C}_n^2$ (Table 1) to be $\sim 25\%$, from the uncertainty in distance to PSR0950+08 and the measurement errors in $\Delta \nu_d$.

The derived values of C_n^2 for PSR0950+08 are 5 to 10 times smaller than those obtained for other pulsars within a few kiloparsec of the Earth⁸. Three of the least strongly scattered pulsars are 1937+21, 2016+28 and 2020+28. These are believed to lie in relatively quiescent regions between the local and Sagittarius spiral arms. Even so, they have values of $\log \hat{C}_n^2$ close to -3.5, an order of magnitude larger than that of PSR0950+08.

There has been considerable progress recently in delineating the galactic distribution of plasma turbulence^{7,8,16,17}. There are at least two components of the 'fluctiferous' medium¹⁷, the 'type A' and 'type B' phases. The A medium is the less strongly scattering of the two. It is characterized by $C_n^2 \approx 10^{-3.5}$ at the galactic midplane, and an exponential z-height (scale height perpendicular to the galactic plane) ~ 1 kpc. The scattering strength of the B medium is about 1,000 times stronger, but it is more important in the inner Galaxy at low galactic latitudes. From models of the type A medium, we would expect $C_n^2 \approx 10^{-3.6}$ for PSR0950+08, about 10 times greater than observed. The type A scattering medium is probably associated with the warm ionized phase of the ISM or with the extended envelopes of H II regions. The local bubble, on the other hand, is related to the coronal phase of the ISM. Theoretical work indicates that the coronal phase cannot support the hydromagnetic waves responsible for galactic scattering^{3,4}, and the low value of C_n^2 inside the local bubble lends observational support to that idea.

Hajivassiliou¹⁶ has argued that the local bubble, although possibly quiescent inside, is surrounded by an envelope of enhanced turbulence, with thickness ~ 0.8 pc and $C_n^2 \approx 0.7 \text{ m}^{-20/3}$. This would explain an observed asymmetry in the sky distribution of extragalactic source sizes. There are at least two possible reasons why we did not observe enhanced scattering from such a shell. First, the distance to the putative plasma shell in the direction of PSR0950+08 is ~ 120 pc, nearly the same as the distance to the pulsar. Scattering screens close to the source do not contribute significantly to the line-of-sight averaged $\hat{C}_n^2$ as inferred from decorrelation bandwidths (see equation 8 in ref. 8). A second possibility is that, in the third quadrant, the bubble is not bounded by a shell and merges gradually into the ambient ISM^{1,18}. If the turbulence responsible for scattering arises upstream of such shells then the weak scattering of PSR0950+08 may be attributable to the absence of a shell or shock wave in this direction.

Although the scattering strength of gas in the local bubble is low, the fractional fluctuation of electron density could be high. For a Kolmogorov turbulence spectrum (equation (1)), the r.m.s. electron density is given by the formula $\delta n_e \approx 0.7 [C_n^2 l_0^2]^{1/2} \text{ cm}^{-3}$, with l_0 in pc and C_n^2 in $\text{m}^{-20/3}$. Assuming

an outer scale $l_0 \approx 100$ pc in the local bubble, we obtain $\delta n_e \approx 0.02 \text{ cm}^{-3}$. From the distance to PSR0950+08 (ref. 10) and its dispersion measure¹⁹, $2.9702 \pm 0.001 \text{ pc cm}^{-3}$, we infer a mean electron density $n_e = 0.023 \text{ cm}^{-3}$. Hence $\delta n_e/n_e \approx 1$. The ratio depends on the assumed value for l_0 , and the adopted value of 100 pc is really only an upper limit based on the size of the bubble. In the diffuse type A scattering medium, an outer scale as small as 0.1 pc is sufficient to make the ratio $\delta n_e/n_e$ equal unity²⁰. If l_0 were of the order of 0.1 pc in the local gas, the fractional fluctuation would be ~ 0.1 , a value comparable to fluctuations in the solar wind near the Earth's bow shock²¹.

Our main conclusion, the smallness of C_n^2 in the local ISM, depends on the assumption of a Kolmogorov turbulence spectrum. Given the ubiquity of the Kolmogorov spectrum for other lines of sight through the ISM, this seems a reasonable assumption. But verification that α is roughly 11/3 is clearly needed. The most common way to estimate α for pulsars is by the frequency scaling of the scintillation bandwidth. This has proved difficult for PSR0950+08 because of the relatively narrow observing bandwidths available at most radio telescopes. This situation may change when wide-band spectrometers become available on the Green Bank telescope and on the upgraded Arecibo telescope. Another approach would be to measure long-term fluctuations in dispersion measure, from which α can be inferred for fluctuation scales between $\sim 10^9$ and 10^{15} cm.

Such measurements are in progress⁶ but may take several years to complete. □

Received 1 June; accepted 14 September 1992.

1. Cox, D. P. & Reynolds, R. J. *A. Rev. Astr. Astrophys.* **25**, 303–344 (1987).
2. McCammon, D. & Sanders, W. T. *A. Rev. Astr. Astrophys.* **28**, 657–688 (1990).
3. Cesarsky, C. J. *A. Rev. Astr. Astrophys.* **18**, 289–320 (1980).
4. Spangler, S. R. *Astrophys. J.* **376**, 540–555 (1991).
5. Rickett, B. J. *A. Rev. Astr. Astrophys.* **28**, 561 (1990).
6. Phillips, J. A. & Wolszczan, A. *Astrophys. J.* **382**, L27 (1991).
7. Cordes, J. M., Spangler, S. R., Weisberg, J. M. & Clifton, T. R. in *Radio Wave Scattering in the Interstellar Medium* (eds Cordes, J. M., Rickett, B. J. & Backer, D. C.) 180–184 (American Institute of Physics, New York, 1988).
8. Cordes, J. M., Weisberg, J. M. & Boriakoff, V. *Astrophys. J.* **288**, 221–247 (1985).
9. Moran, J. M., Rodriguez, L. R., Green, B. & Backer, D. C. *Astrophys. J.* **348**, 147–152 (1989).
10. Gwinn, C. R., Taylor, J. H., Weisberg, J. M. & Rawlings, L. A. *Astr. J.* **91**, 338–342 (1986).
11. Reynolds, R. J. *Astrophys. J.* **348**, 153–160 (1990).
12. Snowden, S. L., Cox, D. P., McCammon, D. & Sanders, W. T. *Astrophys. J.* **354**, 211–219 (1990).
13. Frisch, P. C. & York, D. G. *Astrophys. J.* **271**, L59–63 (1983).
14. Paresce, F. *Astr. J.* **89**, 1022–1037 (1984).
15. Roberts, J. A. & Ables, J. G. *Mon. Not. R. astr. Soc.* **201**, 1119–1138 (1982).
16. Hajivassiliou, C. A. *Nature* **355**, 232–234 (1992).
17. Cordes, J. M., Weisberg, J. M., Frail, D. A., Spangler, S. R. & Ryan, M. *Nature* **354**, 121–124 (1991).
18. Ikeuchi, S., Habe, A. & Tanaka, Y. D. *Mon. Not. R. astr. Soc.* **207**, 909–927 (1984).
19. Phillips, J. A. & Wolszczan, A. *Astrophys. J.* **385**, 273–281 (1992).
20. Cordes, J. M., Wolszczan, A., Dewey, R. J., Blaskiewicz, M. & Stinebring, D. R. *Astrophys. J.* **349**, 245–261 (1990).
21. Spangler, S. R. in *Radio Wave Scattering in the Interstellar medium* (eds Cordes, J. M., Rickett, B. J. & Backer, D. C.) 66–69 (American Institute of Physics, New York, 1988).

ACKNOWLEDGEMENTS. We thank L. Belkora & S. Spangler for important contributions to the arguments in the paper. Arecibo is part of the National Astronomy and Ionosphere Center which is operated by Cornell University under contract with the NSF.

Identification of the nebula G70.7+1.2 as a bow shock powered by a pulsar/Be-star binary

S. R. Kulkarni*, S. N. Vogel†, Z. Wang‡ & D. O. S. Wood§

* Palomar Observatory & Owens Valley Radio Observatory 105-24, California Institute of Technology, Pasadena, California 91125, USA

† Astronomy Program, University of Maryland, College Park, Maryland 20742, USA

‡ Owens Valley Radio Observatory & Infrared Processing and Analysis Center 100-22, California Institute of Technology, Pasadena, California 91125, USA

§ National Radio Astronomy Observatory, Socorro, New Mexico, USA

THE enigmatic radio and optical nebula G70.7+1.2^{1,2} has been suggested to be a young supernova remnant^{3,4}, a nova shell⁴, a cometary⁵ H II region⁶ and a protostellar outflow⁷. Claims^{4,7} of strong nonthermal radio emission cast severe doubts on all but the supernova remnant model, but for this the expansion velocity⁴ would be low. We present here new data that firmly establish the nonthermal nature of the radio emission, and from H α and [O I] Fabry–Perot observations we argue that the extended optical emission arises from a bow shock powered by a mass-losing luminous (Be) star moving supersonically through dense gas. The nonthermal emission is then explained as the shocked relativistic wind from a pulsar, which we propose is a companion to the Be star. The coincidence of the optical and radio emission requires the pulsar and stellar winds to be mixed together. The system has a large overall velocity, $\sim 60 \text{ km s}^{-1}$, which is inexplicable in all other models but which is typical of binary pulsars. Detection of pulsed emission and of the predicted proper motion would confirm our proposal.

The nonthermal nature of G70.7+1.2 is inferred from the detection⁴ of marginal linear polarization and a negative spectral index, $\alpha = -0.5$ (ref. 7); here $S(\nu)$, the flux density at frequency ν , is assumed to be proportional to ν^α . The identification of the

negative spectral index rested entirely on one unpublished measurement at 89 GHz. As the most unusual aspect of G70.7+1.2 is its nonthermal emission we undertook new observations to confirm this. The new measurements (Fig. 1) obtained using telescopes at the Owens Valley Radio Observatory (OVRO) and the Very Large Array (VLA) firmly establish the negative spectral index. Unambiguous proof of synchrotron emission is provided by the detection of 5–10% polarization in the 8.65-GHz image (Fig. 2). These observations rule out all previous models except the supernova remnant model. The low observed velocities⁴ are, however, inconsistent with the supernova remnant model⁷.

At optical wavelengths, extended emission in H α , [S II] and [O I] has been detected^{4,8}. These observations are best fitted with a 100 km s^{-1} shock running into dense (number density n_0) interstellar medium, specifically a small molecular cloud⁷ whose velocity with respect to local standard of rest (v_{LSR}) is $+5 \text{ km s}^{-1}$ and whose distance is 4.5 kpc. An infrared (IR) luminous star with numerous optical emission lines (H α , H β , Fe II and Fe [II]) seems to be embedded in the nebula.

The radio morphology of G70.7+1.2 (Fig. 2) closely resembles known bow shocks⁹. If the usual model is applied to this source, a powerful wind from the IR star inflates a bubble, and stellar motion at velocity v_* with respect to the ambient interstellar medium drives a bow shock. Shocked interstellar medium then produces the bright emission lines. To verify this reasonable hypothesis we undertook 6,300 Å [O I] and H α observations with the Maryland-Caltech Fabry–Perot camera¹⁰ at the Palomar 60-inch telescope.

The [O I] emission also has a bow shock shape coinciding with the nonthermal radio emission (Fig. 2). The emission is broad (full width at half maximum ≈ 70 – 180 km s^{-1}) and mainly blueshifted (0 – 70 km s^{-1}), especially along the symmetry axis. Emission becomes narrower and approaches the ambient gas velocity toward the limb-brightened edges of the nebula, except near the apex where emission is very broad and blueshifted (Fig. 3). These features are predicted by bow shock models¹¹ in which the star is moving toward the observer with an angle $\alpha \approx 60^\circ$ between the velocity vector and line of sight. The inferred stellar speed is $60 < v_* < 120 \text{ km s}^{-1}$. Other models, such as expansion, do not naturally explain the predominance of blueshifts everywhere in the nebula nor the broad, highly blueshifted velocities at the apex. The inferred v_* is less than the 200 km s^{-1}

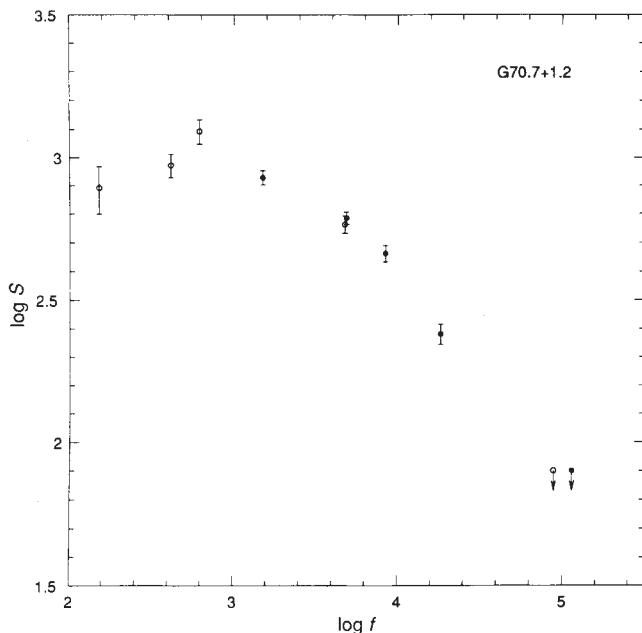


FIG. 1 Radio spectrum of G70.7+1.2. Frequencies (f) are in MHz. O, Flux densities (S , in mJy) taken from the literature; reading left to right, these are from refs 6, 6, 18, 1 and 7; the last refers to an unpublished result quoted in ref. 7. ●, The new OVRO and the VLA measurements. The VLA continuum measurements were obtained in January 1992 with the array in the B/C configuration. The integration time per band was 10 minutes. The VLA data points refer to the flux density within 30 arcsec of the source. The uncertainty is dominated by the limited spatial frequency coverage of this extended (15 arcsec) source. The 115-GHz measurement was obtained with the three-element OVRO interferometer during October 1990 to May 1991. Emission was detected in the OVRO image with a flux density of ~ 50 mJy but it does not resemble the (nonthermal) VLA image at lower frequency. Given the low signal-to-noise ratio and incomplete spatial frequency coverage, we quote an upper limit (A more sensitive OVRO image at 99 GHz by J. A. Phillips (personal communication) shows that the source is clearly detected with a flux density of ≥ 70 mJy.) The 18.5-GHz OVRO measurement was made on 11 March 1992 with the OVRO 40-m telescope and a dual-beam receiving system¹⁹. Some of the measurements in the literature, especially those below a few gigahertz, were made with broad beams and included a contribution from a source 4.5 arcmin northwest of G70.7+1.2, suspected³ to be H II region. Our Fabry-Perot images confirm this to be a distant 1.5-arcmin diameter H II region. We estimate the contribution of this source to be less than 10% at frequencies below a few gigahertz, so its contribution to the measured flux density can be ignored.

broadening seen at the apex (Fig. 3). Similar discrepancies between linewidths and v_* are seen in other bow shocks and have been attributed⁵ to Kelvin-Helmholtz instabilities.

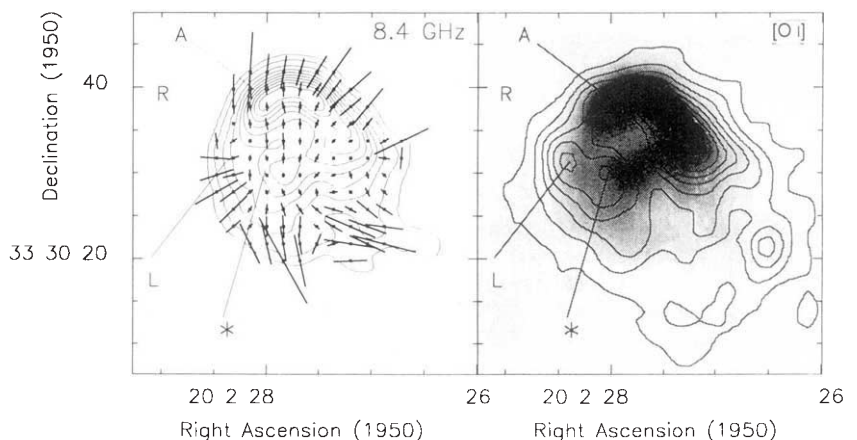
The H α data also support the bow shock model. H α emission is a mix of nebular H α , with profile similar to that of [O I], and H α from the IR star reflected by dust⁸. Comparison with Fabry-Perot 6,450-Å continuum image shows that outside the region of [O I] emission the H α emission is reflected stellar H α . The reflected H α profiles have asymmetries that resemble the stellar profile, except that the lines are broader and typically blueshifted by 10–50 km s⁻¹ (Fig. 3). Again this is expected in the bow shock model and inconsistent with other models. Extinction inferred from our unpublished OVRO carbon monoxide map may distort the morphology in optical lines and explain the low ratio of [O I] to radio emission at the apex. But as the molecular gas is presumably outside the nebula, the optical line profiles are not expected to be affected.

In the framework of the bow shock model, the mean pressure in the nebula must be balanced by the ram pressure, $n_0 m_H v_*^2 \approx 2 \times 10^{-8} n_2 v_2^2$ erg cm⁻³; here m_H is the mass of a hydrogen atom

(gm), and we have defined n_2 and v_2 from $n_0 = 10^2 n_2$ cm⁻³ and $v_* = 10^2 v_2$ km s⁻¹ for ease of handling. Mass loss from the IR star must be inflating the bubble. The absolute J magnitude of the IR star is -4.6 for a visual extinction⁸, $A_v \approx 5.6$. This, the IR peculiar emission features similar to those associated with known early-type stars⁴, the weakness of thermal radio emission, and the optical emission lines⁴ and relatively narrow H α core indicate that the IR star is probably a luminous Be star, with luminosity $L \geq 10^5 L_\odot$ (class II; see ref. 9. $L_\odot$ is the solar luminosity). It is not uncommon for such stars to have powerful winds¹² with mass loss rate $\dot{M} \approx 10^{-5} M_\odot$ yr⁻¹ ($M_\odot$ is the solar mass) and terminal velocity $v_w = 2 \times 10^3$ km s⁻¹. Such a wind provides a nebular pressure $p_{\text{neb}} = \dot{M} v_w / 4\pi r_s^2$ which can balance the ram pressure for $v_2 \leq 1$. Here $r_s \approx 5 \times 10^{17} / \sin(\alpha)$ cm is the standoff distance, the distance between the IR star and the apex of the bow.

The Fabry-Perot data reveal extended emission with large emission measures, $EM T_4^{-0.5} \approx 10^4 - 10^5$ cm⁻⁶ pc; here $T_e = 10^4$ K is the electron temperature and $T_4 \geq 1$ from the H α linewidth, and $EM = \int n_e^2 dl$, the integral of the square of the

FIG. 2 (Left) Emission from G70.7+1.2 at 8.4 GHz mapped with the VLA (contours) at an angular resolution of 3". The orientation of polarized emission at 8.4 GHz is indicated by the vectors. The length of the vectors is proportional to the percentage polarization; the maximum polarization represented is 23%. (Right) Emission 6,300 Å [O I] detected by the Maryland-Caltech Fabry-Perot interferometer (contours) at an angular resolution of 3 arcsec superimposed on the VLA 8.4-GHz image (half-tone). The morphology of the nebula in the radio and in [O I] resembles bow shocks seen around young stars⁵. The [O I] emission (identified with the bow shock in the interstellar medium) coincides closely with the 8.4-GHz emission (identified with the shock of the pulsar wind). The position of the infrared star is indicated by an asterisk and the apex of the bow shock by 'A'. 'R' and 'L' indicate the positions for which spectra, displayed in Fig. 3, of the reflection nebula and the bow shock limb were obtained. The Fabry-Perot data were obtained on the night of 13 March 1992 (UT) with the Fabry-Perot operated at the Cassegrain focus of the Palomar 60-inch telescope. The instrument consists of a re-imaging camera with a Queensgate etalon in the collimated beam. The field of view is 16-arcmin with a pixel size of 1.2 arcsec. The FP data consist of 45 images taken with



uniformly sampled etalon spacings, covering $v_{\text{LSR}} \pm 250$ km s⁻¹ with a spectral resolution (full width at half maximum) of 20 km s⁻¹. The velocity calibration is accurate to 2 km s⁻¹ (r.m.s.).

electron density, n_e^2 , along the line of sight in units of $\text{cm}^{-6} \text{ pc}$. The resulting free-free opacity accounts for the low-frequency turnover at metre wavelengths (Fig. 1). A frequency-independent contribution of $80 T_4 \text{ mJy}$ is expected from the free-free emission over and above the nonthermal contribution of $124(\nu/100 \text{ GHz})^{-0.43} \text{ mJy}$. The millimetre upper limits (Fig. 1) clearly indicate that the nonthermal spectrum must be curved.

The new observations thus strongly support the bow shock model. But none of the known bow shocks around massive stars have detectable nonthermal emission¹³. Nonthermal emission requires energetic particles which can, as in young supernovae and novae, be produced by acceleration in shocks. The shock speed and the wind speed in G70.7+1.2 are, however, considerably lower than in young supernovae. An interpolation of the observed radio emission from novae¹⁴ results in radio emission three orders of magnitude smaller than observed. Bally *et al.*⁷ proposed an 'exotic' model involving an accreting neutron star to power the nonthermal emission. However, no⁴ bright X-ray source is seen coincident with the IR star. The origin of the nonthermal emission from G70.7+1.2 remains mysterious.

We propose that the nonthermal emission is powered by a non-accreting pulsar companion. A wind of relativistic particles and magnetic fields, carrying the pulsar spin-down luminosity, $\dot{E}_R = I\omega\dot{\omega}$ undergoes reverse shock close to the IR star and then expands to fill the nebula; here $\omega = 2\pi/P$ is the angular spin frequency of the pulsar rotating with period P . The minimum energy in magnetic field and particles for the nonthermal radio nebula is estimated to be $5 \times 10^{45} \text{ erg}$. If the pulsar wind travels at c , the speed of light, then the nebula crossing time is 3 yr and the spin-down rate is $\dot{E}_R \approx 3 \times 10^{37} \text{ erg s}^{-1}$, indicating a young ($\lesssim 10^4 \text{ yr}$) pulsar. There is, however, no indication of a young supernova remnant, at any wavelength, radio¹ or optical⁴, suggesting a considerable problem with the proposed model.

The problem can be avoided by assuming that the pulsar wind and the stellar wind act effectively as one fluid with the latter dominating; the pulsar wind is either mixed or (more probably) conected away by the more powerful stellar wind. Thus the effective sound speed is roughly that of the stellar wind v_w , in which case $\dot{E}_R \approx 10^{35} \text{ erg s}^{-1}$. The hypothesized pulsar could either be a middle-aged pulsar ($\lesssim 10^5 \text{ yr}$) slowly rotating (period $P \lesssim 1 \text{ s}$) or an older spun-up pulsar ($P \lesssim 0.05 \text{ s}$). This assumption also naturally explains the close coincidence of the radio

emission with the bow shock in the interstellar medium. The magnetic field inferred from our polarization observations, could either arise, as in plerions, from compression of the ambient field or, more probably, from the hypothesized pulsar. Detailed theoretical work and further polarization observations are needed to clarify this issue.

Such pulsar binaries are predicted by models of massive star evolution¹². Most massive stars begin as binaries. The primary evolves and explodes, leaving a young pulsar in a highly eccentric orbit. Even in the absence of asymmetric stellar collapse, the resulting system acquires velocities comparable to the orbital velocity before explosion. This explains the large inferred v_* , an observation at odds with all other proposed models. The pulsar shines for a while, slows down and in suitable cases, accretes matter from the companion's wind and is spun up (the X-ray binary phase¹²). At some point, the neutron star reappears as a rapidly rotating radio pulsar. Nonthermal particle production is possible in the last two phases. The putative pulsar is unlikely to be millisecond pulsar with low magnetic field ($B \lesssim 10^9 \text{ G}$) because no millisecond pulsar has been found either in the disk or the globular cluster system with a detectable radio nebula; this suggests intrinsic differences in the composition of pulsar wind from pulsars of low and high magnetic field strength.

Several related systems have already been found. A pulsar with a period of 48 ms around a Be star but at an unknown phase of evolution (before or after spin up) has been discovered¹⁵. Following the second supernovae, the spun-up pulsar will be (most probably) flung off from the system, as seems to be the case for the 40-ms pulsar 1951+32¹⁶ in CTB80.

Our proposed model explains the two outstanding features of G70.7+1.2: the nonthermal radio emission and large v_* . The pulsar signal, leaving aside problems of pulsar beaming, should be detectable at sufficiently high frequencies. The copious mass loss assumed here may be directly measurable at radio wavelengths¹⁷, as could the predicted proper motion of 15 mas over a 5-yr period. As well as explaining the origin of this enigmatic source, our model and observations draw attention to two issues of interest to theoretical astrophysics: the interaction of relativistic and stellar winds, and the curvature of the synchrotron spectrum.

Note added in proof: A preliminary search for pulsations using the Arecibo telescope was unsuccessful. Observations at higher frequency, 2.2 GHz, have been scheduled. □

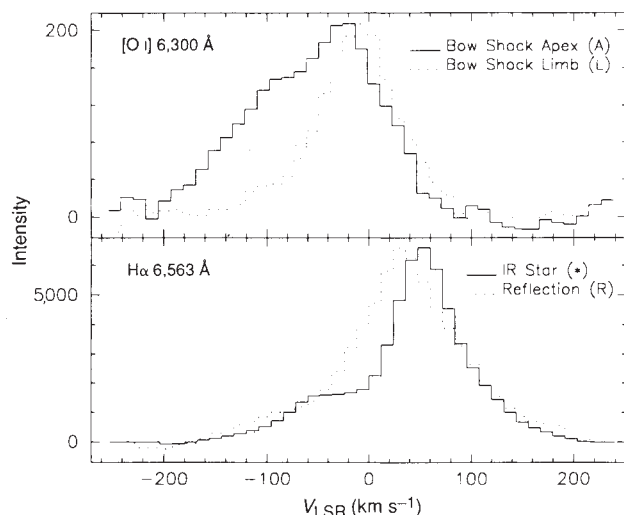


FIG. 3 (Upper panel) Fabry-Perot 6,300 Å [O I] spectrum of bow shock apex (solid line) and limb (dotted line); positions at which spectra were obtained in G70.7+1.2 are marked by A and L, respectively, in Fig. 2. (Lower panel) Fabry-Perot 6,563 Å H α spectrum of infrared star (solid line) and stellar H α emission reflected from dust grains (dashed line); positions at which spectra were obtained are marked in Fig. 2 by the asterisk and R, respectively.

Received 10 June; accepted 14 September 1992.

- Reich, W., Fürst, E., Stuffle, P., Reif, K. & Haslam, C. G. T. *Astr. Astrophys. Suppl.* **58**, 197–248 (1984).
- Green, D. A., *Mon. Not. R. astr. Soc.* **216**, 691–700 (1985).
- Reich, W., Fürst, E., Altenhoff, W. J., Reich, P. & Junkes, N. *Astr. Astrophys.* **151**, L10–L12 (1985).
- de Muizon, M., Strom, R. G., Oort, M. J. A., Claas, J. J. & Braun, R. *Astr. Astrophys.* **193**, 248–264 (1988).
- Van Buren, D., Mac Low, M.-M., Wood, D. O. S. & Churchwell, E. *Astrophys. J.* **353**, 570–578 (1990).
- Green, D. A. *Mon. Not. R. astr. Soc.* **219**, 39–43 (1986).
- Bally, J. *et al. Astrophys. J.* **338**, L65–L68 (1989).
- Becker, R. H. & Fesen, R. A. *Astrophys. J.* **334**, L35–L40 (1988).
- Van Buren, D. & Hester, J. J. *Astrophys. J.* (in the press).
- Vogel, S., Roth, J., Kulkarni, S., Hester, J. & Gruendl, B. *Am. astr. Soc.* **23**, 874 (1991).
- Van Buren, D. & Mac Low, M.-M. *Astrophys. J.* (in the press).
- van den Heuvel, E. J. P. & Rappaport, S. A. *Physics of Be stars* (Sletbak, A. & Snow, T. P.) 291–308 (Cambridge Univ. Press 1987).
- Wood, D. O. S. & Churchwell, E., *Astrophys. J. Suppl. Ser.* **69**, 831–895 (1989).
- Reynolds, S. P. & Chevalier, R. A. *Astrophys. J.* **281**, L33–L35 (1989).
- Johnston, S. *et al. Astrophys. J.* **387**, L37–L42 (1992).
- Kulkarni, S. R. *et al. Nature* **331**, 50–53 (1988).
- Biegling, J. H., Abbott, D. C. & Churchwell, E. B. *Astrophys. J.* **340**, 518–536 (1989).
- Israel, F. P. *Astr. Astrophys.* **48**, 193–212 (1976).
- Readhead, A. C. S. *et al. Astrophys. J.* **346**, 566–587 (1989).

ACKNOWLEDGEMENTS. We thank J. Bally for interesting S.R.K. in this object, T. Herbig and A. Readhead for obtaining the OVRO 22-GHz measurements, G. Vasisht for help with the Fabry-Perot observations, J. A. Phillips for informing us of his results before publication, and R. Blandford, M. Cohen, D. Green, J. A. Phillips and D. Van Buren for discussion. OVRO is supported by a grant from the U.S. NSF. The 60-inch telescope is jointly operated by Caltech and the Carnegie Institution of Washington. The VLA is operated by Associated Universities, Inc., under contract to NSF. We acknowledge financial support from NSF and NASA. S.R.K. thanks the Packard Foundation and the Perkin Fund for support.

Super-absorbency and phase transition of gels in physiological salt solutions

Yong-Qing Zhang, Toyochi Tanaka
& Mitsuhiro Shibayama

Department of Physics and Center for Materials Science and Engineering,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

IONIC gels with the ability to absorb many times their dry weight of water have found widespread use as absorbents in medical, chemical and agricultural applications¹. The dramatic swelling power of these super-absorbent gels results from both the electrostatic repulsion between the charges on the polymer chains, and the osmotic pressure of the counter-ions². In salt solutions such as saline, urine or blood, however, excess Na^+ and Cl^- ions screen the polymer charges and eliminate the osmotic imbalance, effectively changing the properties of the material to that of a non-ionic gel³: this greatly diminishes the swelling power, and hence the utility of these materials under physiological conditions. Here we report the development of a system combining a non-ionic gel with ionized surfactants, which shows super-absorbent behaviour even in the presence of salt. In water, the hydrophobic gel facilitates the formation of spherical surfactant micelles, which mimic the charged sites of an ionic gel. As the salt concentration is increased, the micelles become rod-like, maintaining the electrostatic repulsion along the polymer chains and thereby preserving the swelling power of the gel.

Screening of the long-range electrostatic interaction between neighbouring polymer charge sites by salt ions situated between them may be eliminated either by insulating the charged polymer with a layer of nonpolar medium, or by aligning the charges so closely that the ions cannot fit between them. Here we demonstrate an example of the latter scheme, using combination of a non-ionic hydrophobic gel and an ionized surfactant. This combination has the characteristics of a polyelectrolyte gel^{4,5}. The hydrophobic portion of the surfactant interacts strongly with the polymer chains and forms a thin micellar rod along the polymer chain in the presence of salts⁶.

Ionized hydrophobic gels, such as an acrylic acid/NIPA copolymer gel, undergo a discontinuous volume phase transition between a swollen phase at lower temperatures and a collapsed phase at higher temperatures⁷. The transition is a result of the

competing forces of repulsive electrostatic interaction between the ions on the copolymers acting to swell the gel, and attractive forces between the hydrophobic segments of the copolymers tending to shrink it². A moderately swollen state is also favoured because this allows the maximum possible number of conformations for the gel (this is known as 'rubber elasticity')⁷. The competitive balance among these interactions depends on the temperature and the degree of ionization. The hydrophobic interaction increases with temperature⁷. Water molecules are known to have an ice-like structure around a nonpolar group making the group to be more soluble in water^{5,7}. The water structure melts at higher temperature and the group becomes less soluble and aggregate. Thus the gel shrinks at high temperatures and swells at lower temperatures.

The greater the ionization of the gel, the higher the transition temperature and the larger the volume change at the transition. The volume of the swollen state at low temperatures increases with degree of ionization. When salt is added to the water all these quantities decrease, and at the salt concentrations found in physiological fluids they become similar to those of the non-ionized gel. The swelling power of a gel can be quantitatively represented by the transition temperature and by the volume change at the transition⁸.

Gels were prepared by standard radical polymerization⁷. We dissolved 7.9 g (700 mM) of NIPA (the main polymer constituent), 0.133 g of *N,N'*-methylenebisacrylamide (cross-linker), 240 μl of tetramethylethylenediamine (accelerator) and 40 mg of ammonium persulphate (initiator) in 100 ml of water. The solution was degassed and polymerized in capillaries of 100 μm diameter. The gel was cut into cylinders of length 1 mm which were placed in pure water or NaCl solutions of various concentrations, with and without 2 wt% sodium dodecyl sulphate (SDS). We determined curves for the gel (volume against temperature) for various surfactant concentrations. The phase transition temperature increases with surfactant concentration and saturates at $\sim 2\%$ SDS. We therefore chose this concentration for our experiments. Copolymer gels of 636 mM NIPA and 64 mM sodium acrylate (ionizable group) were prepared as described above for the NIPA gel, to represent the super-absorbent ionic gels currently available.

Figure 1a shows how the volume of the copolymer gel depends on temperature. In pure water the transition temperature was high (83 °C) and the volume change was large (20 times). In 0.15M salt solution, the discontinuous volume transition disappeared and the transition temperature, defined as the inflection point of the swelling curve, was only 43 °C because the swelling power was severely reduced. The volume at 20 °C was seven times smaller than in pure water.

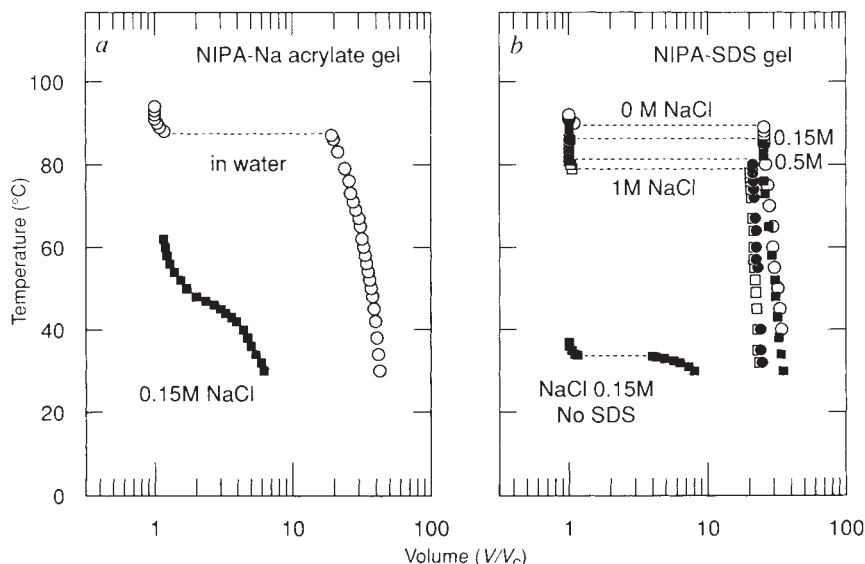


FIG. 1 a, Temperature dependence of the volume of a copolymer gel of poly(*N*-isopropylacrylamide) (NIPA, 636 mM) and acrylic acid (64 mM, ionizable group) in 0.15 mM NaCl solution. V_c denotes the gel volume in the collapsed state at high temperatures. The swelling power is completely lost and the swelling curve becomes the same as that for a non-ionic NIPA gel. b, Temperature dependence of the volume of a non-ionic NIPA gel in NaCl solutions of various concentrations. The solution contains 2 wt% SDS. There is no marked change in the swelling curves, indicating that the gel retains the high swelling power in the salt solutions even at 1M concentration.

The volume of a non-ionized NIPA gel increased threefold at 34 °C (Fig. 1b). When 2% SDS was added to the solution, the transition temperature rose to 90 °C and the volume change became 30-fold. The phase behaviour of the NIPA/SDS gel was not much affected by addition of NaCl even at 1M concentration

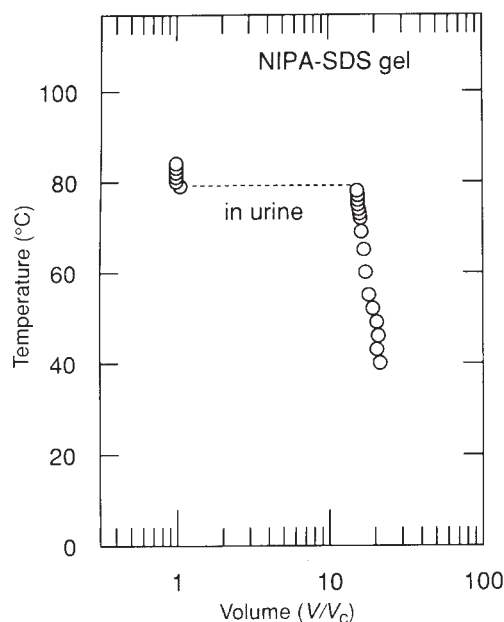


FIG. 2 Temperature dependence of the volume of NIPA gel in human urine. The solution contains 2 wt% SDS. The swelling curve is similar to that for the gel with SDS in pure water. This indicates that the gel retains the high swelling power in human urine.

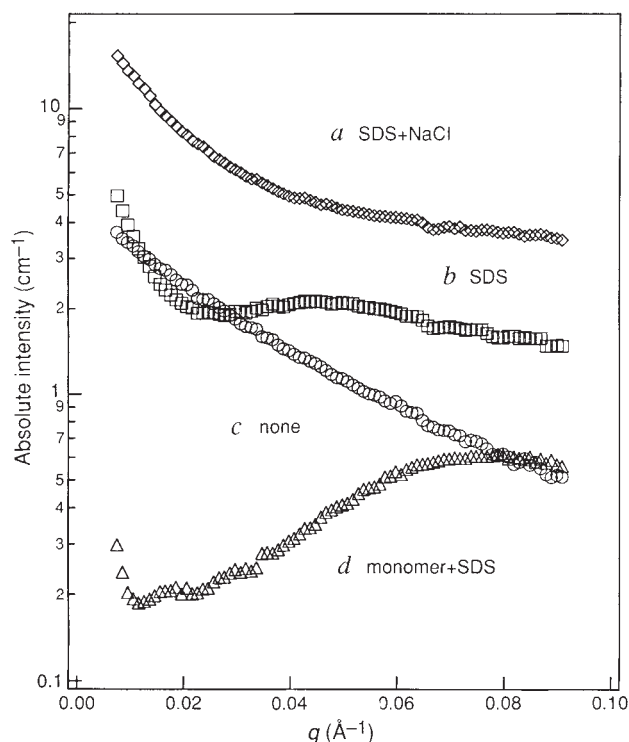


FIG. 3 Absolute intensity of neutrons scattered by the gel-surfactant systems in D₂O, as a function of momentum transfer of scattering. a, NIPA gel with SDS and NaCl; b, NIPA gel with SDS; c, NIPA gel only; d, NIPA monomer with SDS.

(Fig. 1b). Its swelling power was preserved in concentrated salt solutions. Finally, we determined the swelling curve of the NIPA/SDS gel in human urine. Again, no significant change was observed from that in pure water, indicating that the swelling power was indeed preserved under physiological conditions (Fig. 2).

The SDS adsorption to NIPA gel is reversible. It is washed away by water or salt solution. For practical purposes, the gel could be separated from the urine by a thin membrane that would pass water and small ions, but reject surfactant molecules. This membrane would prevent leaching of surfactant from the gel.

To determine the structure of the gel-surfactant system, we did small-angle neutron scattering (SANS) as a function of the momentum transfer, $q = |\mathbf{q}| = (4\pi/\lambda) \sin(\theta/2)$, where θ is the scattering angle and λ is the neutron wavelength. Figure 3 shows the result for (a) NIPA gel with SDS and NaCl, (b) NIPA gel with SDS, (c) NIPA gel only, and (d) NIPA monomer with SDS, all in D₂O. The last of these has a broad peak around $q = 0.07 \text{ \AA}^{-1}$, corresponding to the intermicellar distance, $D = 88 \text{ \AA}$. For the NIPA gel, the scattering curve is a monotonically decreasing function and is characterized by a correlation length $\xi = 24.6 \text{ \AA}$. On addition of SDS to NIPA gel, a scattering maximum appears at $q = 0.045 \text{ \AA}^{-1}$, corresponding to $D = 140 \text{ \AA}$. When NaCl is added, the peak disappears and the scattering intensity increases by a factor of two or more with respect to the results shown in Fig. 3b and c. Our result is consistent with the observation⁹ that SDS forms micelles in water or D₂O and in polyelectrolyte solutions, which give rise to a scattering peak in SANS. As salt is added, the peak shifts towards a high- q region and eventually disappears⁹. Dynamic light scattering showed that the addition of salt transforms the spherical micelles into rod-shaped ones⁶.

From these experimental observations, we propose a structural model for the system. In the absence of salt, the SDS molecules form tight globular micelles, leaving much of the NIPA chains uncovered (Fig. 4). Because of the charge on the

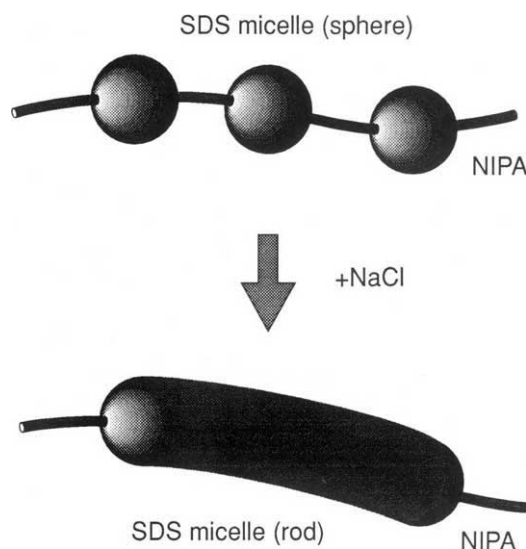


FIG. 4 Illustration of the possible micellar formation of an anionic surfactant in a hydrophobic gel. Ionic surfactants, such as SDS, are known to form spherical micelles in water and rod-like micelles at high salt concentrations⁶. Presence of hydrophobic polymers is expected to aid the micellar formation. When the rod-like micelles cover a polymer chain, because of the closeness between the charges of surfactants in the micelles, the salt ions cannot intercalate between the charges of surfactants in the rod-shaped micelles. Thus the repulsive Coulomb interaction between the neighbouring charges of surfactants is preserved, and so is the swelling power of the gel.

SDS, these micelles repel each other, causing the gel to swell. The gel is practically an ionic one, and the transition temperature is high (90 °C). The micelles, which are separated from each other, give rise to a peak in the neutron scattering curve. On addition of salt, the SDS spreads out along the NIPA polymer, forming cylindrical micelles, and the neutron scattering peak disappears (Fig. 3). In this transformation, as the effective thickness of the polymers increases, so does the scattered intensity.

The model explains why the phase transition behaviour is not affected by the presence of salt. The electrostatic distance becomes smaller than the separation of the charges, as in the conventional super-absorbents in salt solution. In the gel-surfactant system, because the surfactant charge sites in the rod-shaped micelles are so close together, the salt ions cannot be situated between them. The repulsive Coulomb interaction between the neighbouring charges of the surfactants is preserved, and so is the swelling power of the gel.

Gels can thus be designed that are capable of super-absorption and of volume phase transition under physiological conditions.

Such gels will have wide applications in medical and ocean technology as actuators, sensors, controlled delivery systems and super-absorbers. For example as the number of elderly people in our society increases, diapers with a high capacity will be more widely needed. The principle described here will help to improve such technologies. □

Received 18 May; accepted 14 September 1992.

1. Osada, Y. *Adv. Polymer Sci.* **82**, 1 (1987).
2. Tanaka, T. *et al. Phys. Rev. Lett.* **45**, 1636-1639 (1980).
3. Ohmine, I. & Tanaka, T. *J. chem. Phys.* **77**, 5725-5729 (1982).
4. Kokufuta, E., Zhang, Y.-Q., Tanaka, T. & Mamada, A. *Macromolecules* (in the press).
5. Inomata, H., Goto, S. & Saito, S. *Proc. 4th Gel Symp.*, 42-43 (Polymer Society of Japan, Tokyo, 1991).
6. Missel, P. J., Mazer, N., Carey, M. & Benedek, G. B. *Solution Behavior of Surfactants* Vol. 1 (eds Mittel, K. L. & Fendler, E. J.) 373-388 (Plenum, New York, 1982).
7. Hirotsu, S., Hirokawa, Y. & Tanaka, T. *J. chem. Phys.* **79**, 2554-2558 (1982).
8. Ilmain, F., Tanaka, T. & Kokufuta, E. *Nature* **349**, 400-401 (1991).
9. Cabane, B. & Duplessix, R. *J. J. Phys.* **43**, 1529-1542 (1982).

ACKNOWLEDGEMENTS. We thank C. C. Han for the usage of SANS facilities, and W. Robertson, K. Wasserman and K. Tanaka for reading the manuscript. The work was supported by the NSF and Kao Corporation.

Influence of Southern Ocean waters on the cadmium-phosphate properties of the global ocean

Russell D. Frew & Keith A. Hunter

Department of Chemistry, University of Otago, Dunedin, New Zealand

MEASUREMENTS of the cadmium content of marine CaCO_3 deposits, such as aragonitic corals¹ and foraminiferal shells^{2,3}, have been used to trace past ocean circulation² and as an indicator of labile nutrient concentrations⁴. In particular, studies of foraminiferal cadmium have provided insight into the oceanic processes controlling atmospheric CO_2 concentration in glacial times^{5,6}. The use of Cd/Ca ratios in CaCO_3 as a labile nutrient indicator is made possible by a relationship between seawater cadmium and phosphate content that is remarkably uniform throughout the present ocean (Fig. 1). In deep waters of the open ocean, [Cd] at a given value of $[\text{PO}_4^{3-}]$ is consistent within about $\pm 7\%$. However, there is an unusual kink in the relationship at $[\text{PO}_4^{3-}] \approx 1.3 \mu\text{mol kg}^{-1}$ whose cause is not known but which has been suggested² to result from a slightly deeper regeneration cycle for Cd relative to PO_4^{3-} . Points to the right of the kink correspond mainly to intermediate and deep waters of the North Pacific, whereas low PO_4^{3-} concentrations to the left of the kink represent mainly Atlantic and upper-ocean waters⁴. Waters of the Southern Ocean have a strong influence on the composition of the global ocean but few Cd- PO_4^{3-} measurements are available for this region. Here we present recent data for Cd and PO_4^{3-} in Southern Ocean waters, which suggest that the kink is created by the input of Cd-depleted Subantarctic waters into the intermediate waters of the global ocean.

The Southern Ocean is bounded to the north by the subtropical convergence at $\sim 40^\circ\text{S}$, a transition zone between nutrient-depleted, warm surface waters characteristic of mid-latitudes and nutrient-rich, colder waters to the south⁷. The Antarctic convergence ($\sim 55^\circ\text{S}$) divides the Southern Ocean into Subantarctic ($40-55^\circ\text{S}$) and Antarctic regions. The locations of both convergences vary with longitude and season. Antarctic waters are one of the main sources of oceanic deep water, and the Subantarctic region is an important source of water feeding into intermediate depths of the Atlantic and Pacific to the north⁷.

We measured Cd and PO_4^{3-} concentrations in water samples

collected during a cruise of RV *Franklin* on 24 May 1989 at a station located over the Puyseger trench, southwest of New Zealand, at $48^\circ 05' \text{S}$ $164^\circ 30' \text{E}$. Surface temperatures mapped by thermograph and satellite imagery both confirmed that this site was located in Subantarctic waters well south of the subtropical convergence. Water samples were obtained using two independent sampling methods: Teflon-coated 2.5L General Oceanics Go-Flo samplers deployed on Kevlar hydroline with solid polyvinylchloride messengers (0-1,800 m depth), and Teflon-coated 8L General Oceanics Niskin samplers with silicone rubber internal springs and Teflon taps deployed on a conductivity-temperature-depth (CTD) rosette system (0-4,500 m depth). Samples were withdrawn inside a shipboard

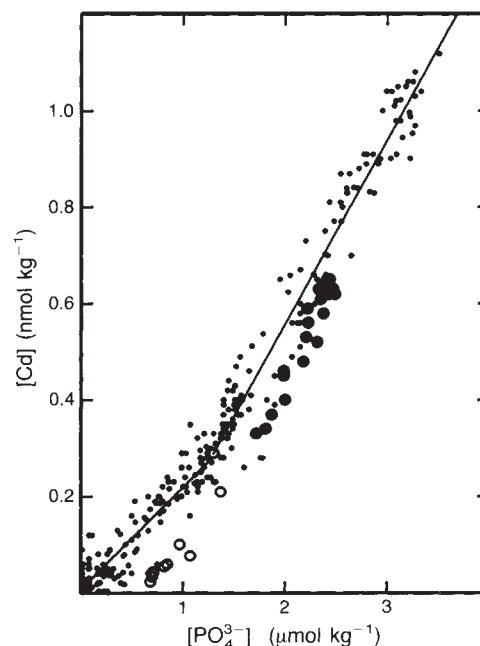


FIG. 1 [Cd] against $[\text{PO}_4]$ for open ocean samples below the mixed layer (●). Data were compiled by E. A. Boyle and comprise those presented in ref. 2 with the addition of Antarctic results (ref. 11). ○, Results obtained in the present study for the Puyseger trench station for depths less than 400 m; ●, depths greater than 400 m. Dotted lines, drawn to emphasize the kink in the global data, are the least-squares regression lines calculated separately for data with $[\text{PO}_4] > 1.3 \mu\text{mol kg}^{-1}$ and for data with $[\text{PO}_4] < 1.3 \mu\text{mol kg}^{-1}$.

TABLE 1 Temperature, salinity, [Cd] and $[PO_4^{3-}]$ for seawater samples

Depth (m)	Temperature (°C)	Salinity	[Cd] (nmol kg ⁻¹)	$[PO_4^{3-}]$ (μmol kg ⁻¹)
25	11.238	34.695	0.041	0.71
25			0.033	0.70
50	11.241	34.697	0.022	0.68
75	11.235	34.700	0.043	0.71
100	11.087	34.770	0.055	0.82
200	10.017	34.754	0.060	0.85
200			0.077	1.07
300	9.077	34.631	0.29	1.3
300			0.10	0.97
400	8.620	34.582	0.21	1.37
700	7.090	34.493	0.34	1.81
750	6.803	34.484	0.37	1.87
790	6.592	34.470	0.33	1.72
900	5.665	34.412	0.40	2.00
990	5.043	34.400	0.45	1.99
1,000	4.942	34.397	0.48	2.18
1,000			0.46	1.99
1,100	4.352	34.417	0.59	2.22
1,190	3.918	34.424	0.53	2.21
1,200	3.922	34.439	0.52	2.32
1,200			0.56	2.23
1,300	3.541	34.471	0.63	2.33
1,500	2.893	34.529	0.62	2.49
1,590	2.765	34.565	0.63	2.38
1,600	2.759	34.566	0.64	2.43
1,600			0.65	2.44
1,750	2.561	34.609	0.63	2.47
2,000	2.222	34.680	0.62	2.43
2,250	2.038	34.705	0.58	2.38
2,500	1.801	34.722	0.61	2.35
3,000	1.507	34.724	0.64	2.37
3,500	1.257	34.719	0.62	2.41
4,000	1.246	34.717	0.65	2.41
4,450	1.263	34.716	0.64	2.41

Samples were collected at the Puysegur trench station, 48° 05' S 164° 30' E. Each row in the table represents a separately collected sample. Each [Cd] entry is the mean of duplicate analyses. Temperature and salinity data were interpolated for each sample depth from CTD data collected using the CTD-rosette system.

clean room into double-bagged polyethylene bottles that had previously been cleaned in acid and filled with an ultrapure acid⁸ cleaning solution. Samples were analysed in duplicate in our clean room laboratory by chelation/solvent extraction and graphite-furnace atomic absorption spectrophotometry⁹. Extraction yield was $92 \pm 4\%$ and procedural blanks were 0.0006 ± 0.0005 nmol kg⁻¹. To check the validity of the results, all samples were independently analysed using a similar technique by D. Mackey of the CSIRO Division of Oceanography, Hobart, Australia. No significant differences were found within experimental error. We measured PO_4^{3-} concentrations on board with a colorimetric autoanalyser system.

The experimental data are presented in Table 1. Temperature-salinity data showed a clear Antarctic Intermediate Water (AAIW) Layer with a salinity minimum of $S = 34.4$ centred on 1,000 m depth, which influenced salinity over most of the upper 2,000 m. The AAIW remains clearly identifiable well north of the Equator in the Atlantic¹⁰. The surface mixed layer was 120 m deep with $S = 34.7$ and $T = 11$ °C. Phosphate concentrations increased with depth (Fig. 2) from 0.7 μmol kg⁻¹ at the surface to a maximum of 2.5 μmol kg⁻¹ below 1,500 m, a value typical of these Subantarctic waters.

The [Cd] profile (Fig. 2) was similar in general shape to $[PO_4^{3-}]$, but showed very low concentrations in the mixed layer (0.02 – 0.03 nmol kg⁻¹). Throughout the mixed layer and most of the AAIW layer (0–1,500 m depth) [Cd] was significantly lower than would be found in mid- and low-altitude water with the same $[PO_4^{3-}]$. When our data are superimposed on the global

Cd-P trend (Fig. 1), the Puysegur trench Subantarctic waters are seen to be significantly depleted in Cd relative to PO_4^{3-} for all but the highest PO_4^{3-} concentrations. This feature was reproduced exactly by results that we obtained at other nearby stations in Subantarctic water.

In Fig. 3 our Subantarctic (48° S) results are compared with data for Antarctic waters in the Drake Passage¹¹ (60° S). Data for latitudes between 48° and 60° S are not at present available. The trend at the two latitudes is consistent, suggesting that the Cd- PO_4^{3-} properties of both Antarctic and Subantarctic waters are influenced by a common process that generates low Cd/ PO_4^{3-} ratios. Although further measurements at latitudes between 48° and 60° S are needed for confirmation, this would imply that a large fraction of the upper water column in Southern Ocean waters has a Cd/ PO_4^{3-} ratio significantly lower than the global trend in Fig. 1.

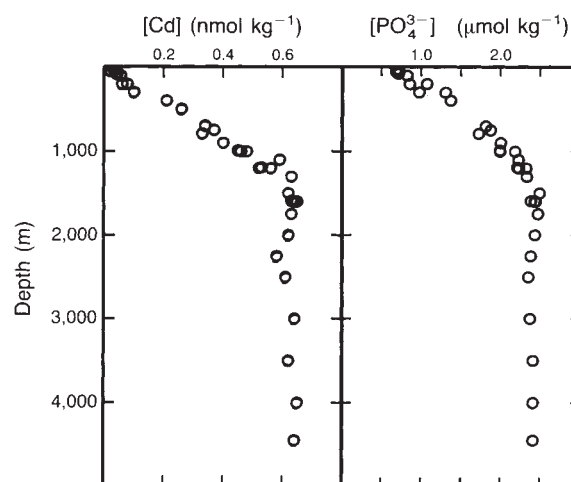


FIG. 2 Vertical profiles of [Cd] and $[PO_4^{3-}]$ for samples collected on 24 May 1989 at a station located over the Puysegur trench southwest of New Zealand at 48° 05' S 164° 30' E in Subantarctic water. Points plotted are averages of duplicate analyses which in all cases, agreed with each other within the experimentally determined precision of $\pm 5\%$.

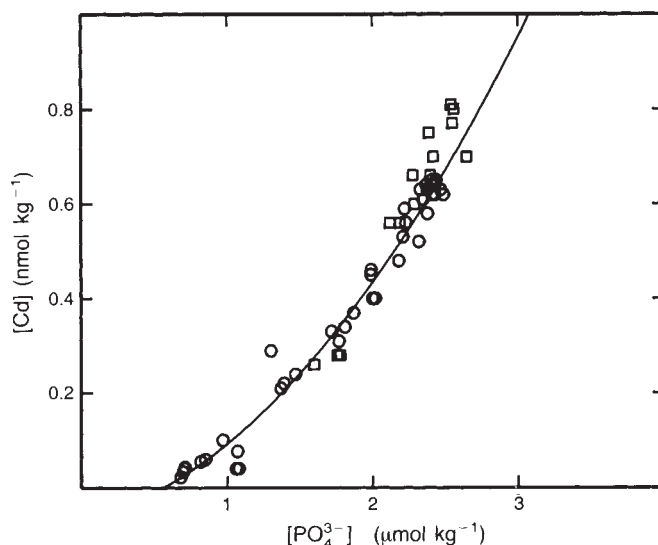


FIG. 3 Cd- PO_4 diagrams for Southern Ocean waters. ○, Puysegur trench station (48° 05' S 164° 30' E, this study), □, Drake Passage (60° 46' S 63° 26' W, ref. 11). A least-squares parabolic regression curve has been plotted through the Puysegur trench results purely to help emphasize the curvature of the data.

The distinct upward curvature of the data in Fig. 3 shows that as biogenic matter containing the phosphorus and cadmium taken up from surface waters sinks down the water column and is remineralized, PO_4^{3-} is regenerated into solution sooner than Cd. Thus the regeneration cycle of Cd is deeper than that of PO_4^{3-} , as suggested earlier². The degree of Cd- PO_4^{3-} curvature in these waters is much greater than that observed in other regions of the ocean, such as those used to compile the global Cd- PO_4^{3-} data in Fig. 1. This implies that in high-nutrient Antarctic and Subantarctic waters, an important fraction of Cd becomes incorporated into components of phytoplankton that are more refractory than the photosynthesized organic tissue that contains the bulk of PO_4^{3-} and other labile nutrients.

One possibility is that Cd becomes incorporated into cellular polyphosphate storage bodies which may form when surplus PO_4^{3-} is available in surface waters¹². Alternatively, the predominance of SiO_2 -secreting organisms (diatoms, radiolaria) in high-nutrient waters may lead to some incorporation of Cd into SiO_2 body parts. By contrast, warm nutrient-poor surface waters are dominated by organisms secreting CaCO_3 . This mineral accumulates Cd in proportion to the seawater Cd/Ca ratio². Per mol of PO_4^{3-} , these phytoplankton take up ~50 mol of Ca (as CaCO_3) and 0.03 mol of Cd¹³. As the Cd/Ca ratio in sea-water is typically less than 10^{-7} , an insignificant amount of Cd ($\sim 50 \times 10^{-7}$) is incorporated into CaCO_3 . The bulk of Cd is presumably taken up in organic tissues in this case.

If Cd were a perfect labile nutrient analogue, the global Cd- PO_4^{3-} trend in Fig. 1 would be a straight line linking the highest Cd and PO_4^{3-} concentrations (North Pacific Deep Water) to the origin, as is the case for correlations between the two nutrients NO_3^- and PO_4^{3-} (ref. 12). In this case, waters in the region of the kink at $1.3 \mu\text{mol kg}^{-1}$ PO_4^{3-} would contain ~0.5 nmol kg^{-1} Cd. In fact, these waters contain considerably lower concentrations of Cd (~0.3 nmol kg^{-1} ; Fig. 1). This reduced [Cd] could be produced by the feeding of Cd-depleted Southern Ocean waters into intermediate depths of the global ocean, provided that intermediate water input is not offset by injection of Antarctic Bottom Water (AABW), whose Cd- PO_4^{3-} properties are nearly 'normal' (Fig. 1). Estimates for the Atlantic¹⁴ suggest a 3:1 ratio of intermediate to bottom water input. A combination of three parts Cd-depleted water having ~0.2 nmol kg^{-1} Cd with one part AABW having ~0.5 nmol kg^{-1} Cd would result in ~0.3 nmol kg^{-1} Cd, a concentration comparable to that found in the region of the kink. More detailed measurements may allow the Cd- PO_4^{3-} deviation from linearity to be used as a measure of intermediate water input from the Southern Ocean.

Clearly the influence of Cd-depleted waters has an important bearing on the use of foraminiferal Cd/Ca ratios as a circulation tracer for past oceans and as a paleonutrient indicator. Circulation studies at intermediate water depths must take account of the Cd content of possible source waters. A single Cd- PO_4^{3-} relationship as in Fig. 1 cannot be used to deduce past nutrient levels in all oceanic regions. Specific relationships must be used in Southern ocean waters, and probably in other high-nutrient areas. Moreover, until we identify the biological mechanism responsible for Cd depletion in Southern Ocean waters, we cannot be certain that the Cd- PO_4^{3-} properties of these waters were the same in the glacial ocean. Changes in circulation or productivity, or the dominance of different phytoplankton species, could each, in principle, have altered the extent and/or geographical location of Cd depletion. □

8. Kuehnen, E. C., Alvarez, R., Paulson, P. J. & Murphy, T. J. *Analyt. Chem.* **44**, 2050-2056 (1972).
9. Danielsson, L. *Mar. Chem.* **8**, 199-215 (1980).
10. McLellan, H. J. *Elements of Physical Oceanography*, 49 (Pergamon, 1965).
11. Martin, J. H., Gordon, R. M. & Fitzwater, S. E. *Nature* **345**, 156-158 (1990).
12. Walsh, R. S. & Hunter, K. A. *Limnol. Oceanogr.* (in the press).
13. Broecker, W. S. & Peng, T.-H. *Tracers in the Sea*, 8-14 (Eldigio, New York, 1982).
14. Tchernia, P. *Descriptive Regional Oceanography*, 161 (Pergamon, Oxford, 1980).

ACKNOWLEDGEMENTS. We thank D. Mackey and E. Butler for assistance with the Franklin samples, and E. Boyle for comments on the manuscript and his Cd- PO_4^{3-} data.

Tilt and northward offset of Cordilleran batholiths resolved using igneous barometry

Jay J. Ague & Mark T. Brandon

Department of Geology and Geophysics, Yale University, PO Box 6666, New Haven, Connecticut 06511, USA

CONSIDERABLE controversy surrounds the suggestion, based on palaeomagnetic evidence, that large segments of the North American Cordillera travelled long distances parallel to the coast during the latest Cretaceous and early Tertiary periods, well after the amalgamation of exotic terranes. Discordant palaeomagnetic data from mid-Cretaceous plutonic rocks of the Peninsular Ranges batholith of coastal southern California and Baja and the Mount Stuart batholith of the Cascade Range in Washington play a pivotal role in the controversy. The discordant data were originally interpreted to reflect northward transport of $\geq 1,000$ km relative to cratonic North America, after the batholiths cooled through their magnetic blocking temperatures¹⁻⁵. More recently it has been argued that the discordances arise from local tilting of batholiths, rather than northward offset^{6,7}. Here we present and implement new methods based on hornblende barometry⁸ for determining the palaeohorizontal in granitic batholiths and correcting palaeomagnetic data for tilting. Our results indicate that the Peninsular Ranges and Mount Stuart batholiths have undergone northward offsets of $\sim 1,000 \pm 450$ and $\sim 2,900 \pm 700$ km, respectively, and also significant tilting.

When a batholith, or any other rock mass, undergoes large-scale displacement or rotation relative to the stable interior of a continental craton, then palaeomagnetic pole locations for rocks of comparable age from the batholith and from the craton will differ. The differences in palaeomagnetic pole locations can be used to quantify post-magnetization latitudinal offset and rotation of the batholith relative to the cratonic reference frame. The palaeomagnetic poles will also be discordant, however, if the batholith has undergone post-magnetization tilting, even if no latitudinal offset or rotation has occurred. It becomes critical, therefore, to assess palaeohorizontal in palaeomagnetic studies, so that measured directions of remanent magnetism can be corrected for the effects of tilting. Interpretation of palaeomagnetic data from calc-alkaline batholiths, such as the Peninsular Ranges batholith (PRB) and the Mount Stuart batholith (MSB), shown in Fig. 1 has long been hampered by the fact that granitic rocks preserve no internal stratigraphy from which palaeohorizontal can be reliably determined.

Our new method for addressing the problem of pluton palaeohorizontal depends on estimates of magma crystallization depth. We obtain these estimates using igneous hornblende barometry⁸. Two empirical^{8,9} and three experimental¹⁰⁻¹² studies have shown that, in the presence of the appropriate buffer assemblage, the total aluminum content of hornblende in a crystallizing granitic magma is a sensitive linear function of pressure. Although there are differences between published barometer calibrations⁸⁻¹², they all have similar slope coefficients. Our results, therefore, are largely independent of the calibration used because we are only concerned with the

Received 10 June; accepted 8 September 1992.

1. Shen, G. T., Boyle, E. A. & Lea, D. W. *Nature* **328**, 794-795 (1987).
2. Boyle, E. A. *Paleoceanography* **3**, 471-489 (1988).
3. Boyle, E. A. *J. geophys. Res.* **93**, 15701-15715 (1988).
4. Boyle, E. A. & Keigwin, L. D. *Nature* **330**, 35-40 (1987).
5. Boyle, E. A. *Geochim. cosmochim. Acta* **50**, 265-276 (1986).
6. Broecker, W. S. & Peng, T.-H. *Global biogeochem. Cycles* **3**, 215-239 (1989).
7. Tchernia, P. *Descriptive Regional Oceanography*, 43-45 (Pergamon, Oxford, 1980).

TABLE 1 Sample locations and hornblende aluminium content

Batholith	Sample	x (km) (+East)	y (km) (+North)	Z _E (km)	Al ^T
PRB (West)	1011-191	-43.12	99.81	0.31	1.31
	1011-195R	-22.74	86.87	0.49	1.35
	1011-196	-13.88	87.80	0.46	1.61
	1011-201	-23.27	28.19	0.12	1.09
	1011-211	4.26	32.35	0.31	1.58
	1011-212	3.88	29.57	0.31	1.61
	1011-213	3.10	25.88	0.46	1.62
	1011-214	-1.94	25.42	0.40	1.32
	1011-238R	32.64	9.70	0.73	1.90
	1011-247	13.22	6.47	0.40	1.45
	1011-249	5.44	10.17	0.21	1.47
	1011-250	0.77	10.17	0.12	1.17
	SC-69-1b	-48.23	94.60	0.24	0.88
	SC-69-36	-10.18	84.22	0.55	1.62
	SC-69-39	6.26	67.77	0.79	1.65
PRB (East)	1011-224	25.08	79.02	1.46	1.59
	1011-225	25.84	83.18	1.68	1.66
	1011-227	16.18	93.80	1.40	1.57
	1011-228	43.29	63.31	0.79	1.69
	1011-230	51.77	64.74	1.22	1.96
	1011-231	53.94	67.00	1.07	2.05
	1011-232	55.20	69.25	0.92	2.08
	1011-234	21.65	60.53	1.13	1.31
	1011-236	17.42	44.82	0.82	1.38
	1011-239	46.24	11.29	0.73	1.74
	1011-241	53.54	25.88	0.67	1.77
	1011-242	52.77	25.42	0.79	1.94
	1011-243	50.83	24.03	0.98	1.79
	1011-244	47.34	23.57	1.28	1.59
	SC-69-40	21.76	78.06	1.07	1.76
	SC-69-41	26.06	77.72	1.40	1.77
	SC-69-43A	37.54	62.90	1.46	1.82
	SC-69-46	52.30	64.66	1.16	1.98
	SC-69-47	65.02	79.42	0.07	2.25
	SC-69-153	72.74	63.00	0.00	2.23
MSB	MS-28	9.70	-13.49	1.54	0.87
	MS-37	12.11	-13.06	1.53	0.83
	MS-47	11.91	-14.42	1.89	0.86
	MS-53	22.05	-8.84	0.67	0.92
	MS-86	16.42	-15.14	2.11	0.89
	MS-94	13.84	-15.21	2.38	0.85
	MS-101B	24.26	-3.30	0.37	0.95
	MS-103	24.35	-3.24	0.37	1.00
	MS-110B	24.23	-3.26	0.37	0.89
	MS-112	-1.88	0.19	1.18	1.20
	MS-120	-2.00	1.52	1.70	0.95
	MS-131	-3.77	-2.50	1.23	1.02
	MS-134	-4.63	-2.71	1.59	1.10
	MS-140	-4.94	-2.39	1.55	0.94
	MS-147	-14.29	9.41	0.77	1.13
	MS-149	-14.26	9.14	0.82	1.16
	MS-151	-13.98	8.41	0.88	1.22
	MS-152	-13.95	8.30	0.91	1.13
	MS-156	-13.36	6.12	0.93	1.16
	MS-171	-13.07	15.34	1.04	1.26
	MS-179	-15.21	11.03	0.50	1.36
	MS-194	-9.79	11.17	0.78	1.16
	MS-199	-9.28	11.16	0.85	1.23
	WP-454	6.21	-15.58	2.13	0.87
	WP-545	6.27	-6.37	1.87	1.14
	WP-549	5.60	-8.34	2.07	0.81

For PRB and MSB, $x=0$ km, $y=0$ km at 243° E, 33° N and 239° E, 47.617° N, respectively. Z_E , sample elevation with respect to sea level. Al^T , mean total Al formula units in rims of euhedral hornblende crystals (23 oxygen basis; all Fe^{2+}) coexisting with critical assemblage required for barometry⁸⁻¹². Setting Fe^{3+}/Fe^{2+} to zero has a negligible effect on estimated crystallization pressures (≈ 0.1 kbar). Sample standard deviation in Al^T is less than 0.14 for all specimens. Prefix 1011-, SC-69-, and WP-samples from refs 14, 15 and 17, respectively.

gradient in palaeopressures, not the absolute accuracy of each pressure determination. We calculate crystallization depth from estimated pressures by assuming an average crustal density of 2.8 g cm^{-3} .

To determine the orientation of palaeohorizontal, the least-squares method is used to fit x - y sample location data and crystallization depth estimates to a planar basis function:

$$Z_{TOT} = a_0 + a_1x + a_2y \quad (1)$$

where Z_{TOT} is the vertical distance from present sea level to the land surface at the time of crystallization (palaeosurface), and a_i are fit parameters. The key assumption here is that the palaeosurface was on the average parallel to palaeohorizontal, an assumption that is compatible with the physiography of analogous modern batholithic provinces, such as the Andes¹³. Z_{TOT} is then related to the depth of crystallization inferred from barometry, Z_B , by

$$Z_{TOT} = Z_B / \cos \beta + Z_E \quad (2)$$

where β is the palaeosurface dip relative to present horizontal, and Z_E denotes sample elevation (Fig. 2). This nonlinear least-squares problem is solved iteratively using standard methods.

Testing the tilt and translation hypotheses requires barometric and palaeomagnetic data for the plutons and palaeomagnetic data for the cratonal reference frame. Previously published PRB barometric data¹⁴ are augmented here with new analyses of appropriate samples from ref. 15 (Table 1). Some reconnaissance barometric results exist for the MSB¹⁶, but we use new determinations obtained specifically for the purpose of constraining palaeohorizontal. The bulk of the MSB samples were collected by us, but several samples from Pongsapich¹⁷ have also been studied (Table 1). Published palaeomagnetic data for the PRB and MSB are from the western and southwestern portions of the batholiths, respectively¹⁻³ (Table 2). Because these parts of the batholiths crystallized at relatively shallow crustal levels, they were probably fully magnetized before tilting¹⁸. Plutons in the western PRB crystallized mainly between 120 and 100 Myr (ref. 19), whereas the MSB has an igneous age of 93-96 Myr (refs 20, 21). The cratonal reference frame for this period is well determined from widely separated sampling sites in North America; virtual geomagnetic pole data used here are from refs 22, 23 (Table 2).

Uncertainties in our estimates of palaeohorizontal orientation, terrane displacement and terrane rotation are best analysed using non-parametric bootstrap statistical techniques^{24,25}, which allow us to assess fully the random errors associated with the barometric data from the pluton, and the palaeomagnetic data from the pluton and the cratonal reference. These methods entail repeated resampling of the original data sets with replacement in order to form a large number of hypothetical data sets, or 'bootstrap samples'. The means calculated from each of these bootstrap samples, referred to as 'replicates', are then used to determine the standard error and confidence region of an estimated value²⁴, such as the mean dip angle of palaeohorizontal. The replicates are calculated according to the following three steps. (1) The planar regression model is used in conjunction with bootstrap resampling of residuals²⁴ to determine a replicate palaeohorizontal tilt by least squares. (2) The mean magnetization direction (declination and inclination) measured at each sampling site in the pluton is corrected according to this replicate palaeohorizontal tilt. These corrected palaeomagnetic directions are resampled, and a replicate palaeomagnetic pole for the pluton is computed. (3) The published virtual geomagnetic pole locations for sampling sites comprising the cratonal reference data set are resampled to derive a replicate palaeomagnetic pole for the craton. These steps are repeated $\sim 20,000$ times to construct distributions of replicate results that map out the mean palaeomagnetic pole for the tilt-corrected pluton and the cratonal reference. These two distributions are used, in turn, to estimate the observed and expected inclination, declination and

TABLE 2 Tilt, northward offset and rotation of the PRB and MSB

Schmidt ¹² calibration of hornblende barometer							
Batholith	Palaeohorizontal Strike/Dip (deg)	Observed Plat (deg)	Observed Dec (deg)	Expected Plat (deg)	Expected Dec (deg)	Northwards Offset (km)	Rotation (deg)
PRB (West)	160 ± 5.5/19 ± 5.4 W	34.8 ± 3.3	342.5 ± 8.4	43.7 ± 2.1	342.2 ± 3.3	990 ± 440	0 ± 9
PRB (East)	161 ± 7.9/15 ± 3.3 W						
MSB	55 ± 38.1/8 ± 3.0 SE	32.0 ± 6.1	16.3 ± 8.4	58.1 ± 2.2	336.9 ± 4.5	2,900 ± 720	-39 ± 10
Johnson and Rutherford ¹⁰ calibration of hornblende barometer							
PRB (West)	160 ± 5.5/17 ± 4.8 W	34.7 ± 3.4	345.3 ± 7.4	43.7 ± 2.1	342.2 ± 3.3	1,000 ± 440	-3 ± 8
PRB (East)	161 ± 8.0/13 ± 2.9 W						
MSB	55 ± 39.9/7 ± 2.5 SE	31.0 ± 5.4	15.5 ± 8.2	58.1 ± 2.2	336.9 ± 4.5	3,020 ± 650	-39 ± 9

All uncertainties are 95% confidence limits. Observed Plat and Dec, tilt-corrected palaeolatitude and declination for the pluton. Expected Plat and Dec, expected palaeolatitude and declination for the pluton. Rotations are positive counterclockwise. Uncertainty estimates calculated using 20,000 bootstrap replications. The small downward bias inherent in bootstrap standard deviation estimates was removed by selecting $n-1$ observations for each bootstrap sample, where n is the number of observations in the original data set of interest²⁴. The pluton palaeomagnetic data comprise 32 and 17 site-mean directions for the PRB and MSB, respectively¹⁻³. Vertical geomagnetic pole locations for cratonic North America (85 sampling sites) are from the 124 Myr Monterege Hills intrusions in Quebec, Canada²² and the 100 Myr Magnet Cove intrusions in Arkansas, USA²³.

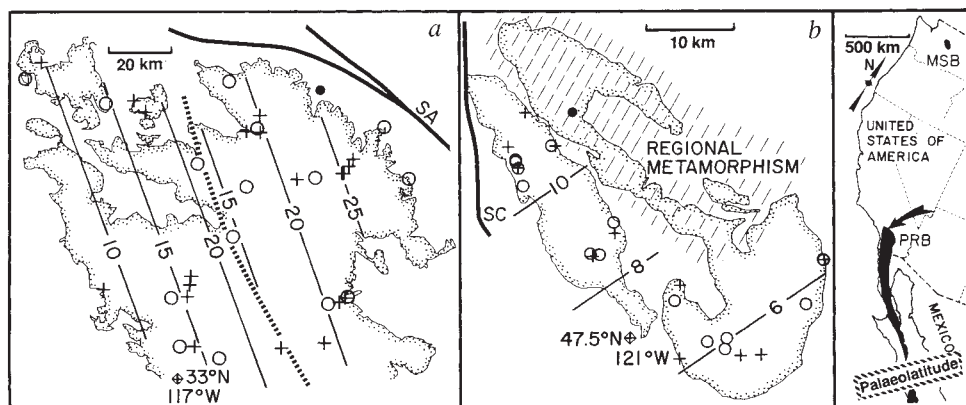
palaeolatitude for the pluton, from which the mean amounts of northward offset and rotation, and associated confidence limits, are computed. The observed values reflect the location and attitude of the batholith when it was magnetized, whereas the expected values are those that would be expected for the batholith if it had always been fixed to the craton. Our bootstrap methods have been tested against standard palaeomagnetic statistical methods for comparable cases where no tilt correction is required²⁶. We have found that the bootstrap-estimated means and confidence intervals are in excellent agreement (within 0.1°) with those computed using standard methods.

Our palaeohorizontal determination and error analysis methods can be used with any type of geobarometric data, either

from plutons or their intruded wallrocks. Inferring patterns of batholith tilting from wallrock aureole barometry, however, is subject to considerable uncertainties associated with the timing of contact metamorphism and the relative movement between magma and wallrock which occurs during intrusion.

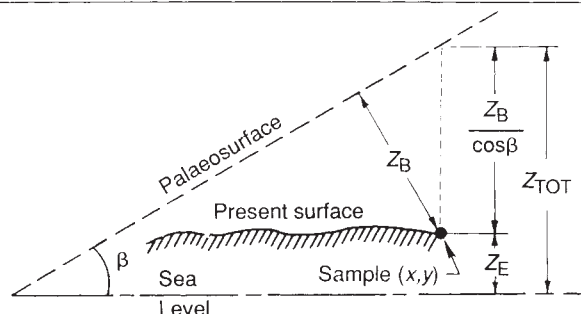
Analysis of pluton barometric data indicates that palaeosurfaces for the PRB and MSB are not parallel to present horizontal (Fig. 1a and b; Table 2a and b). For comparison purposes, we calculate palaeohorizontal tilt using the two experimental hornblende barometer calibrations relevant for the PRB and MSB^{10,12}. The two calibrations yield nearly identical results. Our barometric data reveal a major fault with 7 km of dip separation cutting through the central PRB (Fig. 1a). This fault correlates

FIG. 1 Contour maps of height (Z_{TOT} ; km) from present sea level to best-fit palaeosurface for the batholiths computed using hornblende barometer calibration of ref. 12 (see text and Fig. 2). +, O, Samples having residuals above and below the best-fit surface, respectively. Average of absolute values of residuals is 1.2 km; maximum residual is 3.74 km. Location of batholiths shown in black on inset location map on right. a, Northern Peninsular Ranges batholith (PRB), California, USA (indicated by arrow on inset location map). SA, San Andreas fault system; ●, Palm Springs. Dashed line denotes major fault in the central part of the batholith (see text). Attitudes of best-fit palaeosurfaces to the west and east of this fault are nearly identical (Table 2). b, Mount Stuart batholith (MSB), Washington, USA. SC, Straight Creek fault; ●, Stevens Pass. The northeastern MSB has probably been overprinted by a regional metamorphic event²⁹ (dashed lines). Palaeomagnetic and barometric data used here are from the southwestern MSB. The southwestern MSB and its contact aureole show no evidence of metamorphic overprinting following intrusion. Backscattered electron imaging in the electron microprobe indicates that



textures reflecting subsolidus re-equilibration of hornblende, including exsolution and 'patchy' recrystallization textures³⁰, are absent from all analysed MSB specimens. The excellent fit of the barometric data to the planar palaeosurface model indicates that the western and eastern PRB and southwestern MSB were tilted as coherent crustal blocks. Diagonal ruled line at bottom of location map (right) indicates approximate palaeolatitude, relative to present-day geography, of the MSB and northern PRB before northward offset (see text).

FIG. 2 Geometry of the least-squares model embodied in equation (2) of the text. View is perpendicular to the strike of the palaeosurface. For a given sample located at map coordinates (x, y), Z_B is the depth of crystallization estimated from barometry, Z_E is the sample elevation, β is the palaeosurface dip relative to present sea level, and Z_{TOT} is the vertical distance from present sea level to the palaeosurface.



with several petrologic transitions, including the well known step in oxygen isotope ratios^{19,27,28}. Tilt correction of the western PRB palaeomagnetic data leads to the conclusion that the batholith has been translated northward $\sim 1,000 \pm 450$ km (95% confidence) with no significant rotation since its crystallization (Fig. 1; Table 2a and b). Only about 300 km of this northward translation can be attributed to Neogene opening of the Gulf of California⁷. The MSB crystallized at shallower crustal levels than the bulk of the PRB (Fig. 1b). Our analysis indicates that the MSB has been translated $\sim 2,900 \pm 700$ km to the north, and rotated clockwise by $\sim 40 \pm 10^\circ$ since the mid-Cretaceous (Fig. 1; Table 2a and b).

Our results strongly suggest that the discordant palaeomagnetic data from the PRB and MSB are the result of tectonic tilting, northward translation, and, in the case of the MSB, clockwise rotation, following mid-Cretaceous intrusion of the batholiths. Our tilt corrections do not greatly change the large northward offsets estimated in the original palaeomagnetic studies of the batholiths, but they do lead to much smaller rotation estimates for the PRB^{1,3}. It is interesting that the calculated palaeolatitudes for the PRB and MSB are indistinguishable (Fig. 1; Table 2a and b), suggesting that they originated from the same part of the Cretaceous Cordilleran magmatic arc. □

Received 27 May; accepted 5 October 1992.

1. Teissere, R. F. & Beck, M. E. *Jr Earth planet Sci. Lett.* **18**, 296–300 (1973).
2. Beck, M. E. *Jr*, Burmester, R. F. & Schoonover, R. *Earth planet. Sci. Lett.* **56**, 336–342 (1981).

3. Hagstrum, J. T., McWilliams, M., Howell, D. G. & Grommé, S. *Geol. Soc. Am. Bull.* **96**, 1077–1090 (1985).
4. Irving, E., Woodsworth, G. J., Wynne, P. J. & Morrison, A. *Can. J. Earth Sci.* **22**, 584–598 (1985).
5. Umhoefer, P. J. *Tectonics* **6**, 377–394 (1987).
6. Butler, R. F., Gehrels, G. E., McClelland, W. C., May, S. R. & Klepach, D. *Geology* **17**, 691–694 (1989).
7. Butler, R. F., Dickinson, W. R. & Gehrels, G. E. *Tectonics* **10**, 561–576 (1991).
8. Hammarstrom, J. M. & Zen, E.-an *Am. Miner.* **71**, 1297–1313 (1986).
9. Hollister, L. S., Grissom, G. C., Peters, E. K., Stowell, H. H. & Sisson, V. B. *Am. Miner.* **72**, 231–239 (1987).
10. Johnson, M. C. & Rutherford, M. J. *Geology* **17**, 837–841 (1989).
11. Thomas, W. M. & Ernst, W. G. in *Fluid–Mineral Interactions: A Tribute to H. P. Eugster* (Geochemical Society, 1990).
12. Schmidt, M. W. *Contrib. Miner. Petrol.* **110**, 304–310 (1992).
13. Grow, J. A. & Bowin, C. O. *J. geophys. Res.* **80**, 1449–1458 (1975).
14. Ague, J. J. & Brimhall, G. H. *Geol. Soc. Am. Bull.* **100**, 912–927 (1988).
15. Armstrong, R. L. & Suppe, J. *Geol. Soc. Am. Bull.* **84**, 1375–1392 (1973).
16. Anderson, J. L. & Paterson, S. R. *Geol. Soc. Am. Abstr. Prog.* **23**, A387 (1991).
17. Pongsapich, W. thesis, Univ. Washington (1974).
18. Beck, M. E. *Jr* *Tectonics* **11**, 297–302 (1992).
19. Silver, L. T., Taylor, H. P. Jr & Chappell, B. W. in *Mesozoic Crystalline Rocks–Peninsular Ranges Batholith and Pegmatites, Point Sal Ophiolite*, Guidebook Geol. Soc. Am. Ann. Mtg (1979).
20. Tabor, R. W. et al. *Geologic Map of the Chelan 30-minute by 60-minute Quadrangle, Washington* (Miscellaneous Investigations Series Map I-1661, US Geological Survey, 1987).
21. Walker, N. W. & Brown, E. H. *Geology* **19**, 714–717 (1991).
22. Foster, J. & Symmons, D. T. A. *Can. J. Earth Sci.* **16**, 1716–1725 (1979).
23. Globberman, B. R. & Irving, E. *J. geophys. Res.* **93**, 11721–11733 (1988).
24. Efron, B. *The Jackknife, the Bootstrap, and other Resampling Plans* (SIAM, Philadelphia, 1982).
25. Tauxe, L., Kylstra, N. & Constable, C. *J. geophys. Res.* **96**, 11, 723–740 (1991).
26. Demarest, H. H. Jr *J. geophys. Res.* **88**, 4321–4328 (1983).
27. Taylor, H. P. Jr *Trans. R. Soc. Edinb. Earth Sci.* **79**, 317–338 (1988).
28. Todd, V. R., Erskine, B. G. & Morton, D. M. in *Metamorphism and Crustal Evolution of the Western United States* (Prentice-Hall, Englewood Cliffs, 1988).
29. Evans, B. W. & Berti, J. W. *Geology* **14**, 695–698 (1986).
30. Blundy, J. D. & Holland, T. J. B. *Contrib. Miner. Petrol.* **111**, 269–272 (1992).

ACKNOWLEDGEMENTS. We thank J. L. Anderson, M. E. Beck Jr, R. F. Butler, G. E. Gehrels, E. Irving, R. B. Miller, S. R. Patterson and E.-an Zen for comments and discussion, J. D. Blundy and L. Tauxe for reviews, and R. Burger, D. A. Evans and G. E. Tenzer for assistance in the field. This work was supported by the NSF.

Deep origin of mid-ocean-ridge seismic velocity anomalies

Wei-jia Su, Robert L. Woodward
& Adam M. Dziewonski

Department of Earth and Planetary Sciences, Harvard University,
Cambridge, Massachusetts 02138, USA

ZHANG and Tanimoto¹ have used the results of a new high-resolution three-dimensional model of shear-wave velocities in the Earth's upper mantle² to conclude that, whereas hotspots are underlain by low-velocity anomalies extending to ~ 200 km depth, the corresponding anomalies under mid-ocean ridges are shallower (~ 100 km) in extent. Here we show that such a shallow origin for the mid-ocean-ridge anomalies is inconsistent, by a factor of 3–5, with the observed travel-time residuals for the seismic phase SS (ref. 3). A three-dimensional model describing shear velocities in the entire mantle⁴, derived independently of the SS observations but based on a large set of waveforms (15,000 seismograms⁵) and differential travel-time data (8,200 measurements^{6,7}), shows the velocity anomalies associated with all mid-ocean ridges as continuous features down to 300 km depth. Some of these low-velocity anomalies continue throughout the upper mantle, and may extend into the lower mantle.

Visual comparison of Zhang and Tanimoto's model¹ (ZT) with, for example, M84C (ref. 8; an early three-dimensional model of the upper mantle) reveals differences other than those implied by the nominal resolution of both models: maximum harmonic degree 36 in ZT compared with 8 in M84C. The velocity anomalies in ZT, which is defined to a depth of 500 km, are smaller just below the Moho and decrease with depth rapidly to become insignificant below 300 km depth, whereas variations of $\pm 1.5\%$ in M84C in the transition zone, present also in model MDLSH of Tanimoto⁹, are consistent with the 'degree-2' anomaly of Masters *et al.*¹⁰

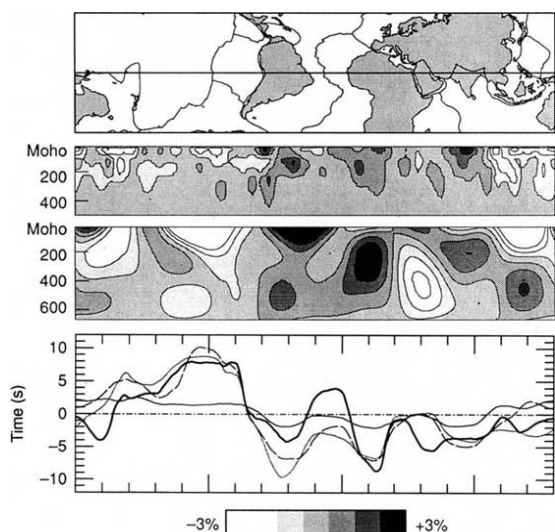
Clearly, both models cannot be right, and each has different geodynamic implications. Model ZT, which is further discussed

in refs 11 and 12, gives an impression that the heterogeneity is concentrated within the top 200–300 km, with the model becoming radially symmetric for practical purposes at greater depths. Other studies^{4–10,13–16} indicate considerable lateral heterogeneity through the upper and lower mantle, although there is general agreement that the root-mean-square (r.m.s.) heterogeneity near the surface is the highest. With models of the latter sort, tomographic modelling of geodynamic observables, in other words the geoid, topography core–mantle boundary and plate motions, has been rather successful^{13,17–20}. It is important to test ZT and one of the earlier models, for example M84C, preferably with data not used in the derivation of either model.

Global observations of the cap-averaged³ (see Fig. 1d) travel times of the phase SS, a downward radiated shear wave reflected once at the Earth's surface, are convenient for such tests. The path through the upper mantle, although not purely vertical, is steep. The sensitivity of the SS residuals to the structure near the surface and, say, in the transition zone is comparable, but for surface waves the sensitivity to the shallow structure can be higher by an order of magnitude. Both ZT and M84C were derived using long-period surface waves.

The great circle path in Fig. 1a intersects the East Pacific Rise (EPR) at roughly the same place as in Fig. 1 of ref. 1. Figure 1b shows the cross-section along this great circle through model ZT, Fig. 1c the cross-section for M84C. Despite the differences in lateral resolution and amplitudes of the anomalies, the two models are somewhat similar within the top 100 km, but then they begin to diverge. In particular, in M84C the EPR anomaly persists to 300 km, and perhaps deeper, but in ZT it is limited to the top 100 km. The amplitudes in ZT fall to the background level below 300 km.

Figure 1d shows that the SS anomalies predicted by model ZT are limited to ± 2 s and are smaller than observations by at least a factor of 3. The anomalies predicted by M84C, except for an offset between longitudes 60° W and 0° , are of roughly the size observed, reaching $+10$ s at the EPR and -6 s under the Sahara. The fact that the observed residuals correlate well with the surface tectonic features means that most of the SS anomaly is accumulated at shallow depths.



The conclusion is that there is considerable heterogeneity throughout the upper mantle. The observed SS anomaly associated with the East Pacific Rise corresponds to a one-way travel time of 5 s. If we assume that the maximum anomaly is constant at -5% , then, with the average velocity of 4.5 km s^{-1} , we need 450 km of ray path to accumulate a 5-s residual. As the ray path is not purely vertical, 350 km thickness might be enough, but even this is significantly greater than the 100 km proposed in ref. 1. To confine the anomaly to the top 100 km one would need -20% heterogeneity over a relatively wide ($\sim 1,000 \text{ km}$) area on both sides of the ridge axis.

The fact that ZT underestimates the amplitude and depth extent of the mid-ocean-ridge anomalies is confirmed by a series

FIG. 1 Mantle structure and SS travel-time anomalies. *a*, Map showing the great circle path (heavy line in the centre) along which the cross-section is made. *b*, Cross-section through model of Zhang and Tanimoto¹. The top of the panel corresponds to the depth of the Moho, the bottom to 500 km. The scale bar gives the velocity anomalies in per cent; near the surface there is some saturation of the scale. *c*, Same as *b*, but for model M84C (ref. 8); bottom of the panel is at 670 km depth. There is considerable saturation: the maximum anomaly at the top is -5% . *d*, Observed and calculated SS residuals averaged over 15° spherical caps. By averaging the anomalies measured between different sources and receivers but having the mid-path reflection point within the same circle of a chosen radius, we reduce the effect of errors in individual measurements and partly remove the effect of path differences away from the nearly common reflection point. SS residuals processed in such a way are primarily sensitive to the structure near the centre of the cap. Heavy line, observed SS residuals from ref. 3. The SS residuals have been corrected for the thickness of the crust and bathymetry or topography. Thin solid line, residuals predicted by model ZT (ref. 1); dashed line, model M84C (ref. 8); dotted line, model SH8/WM13 (ref. 4). The vertical scale is in seconds; a division on the horizontal scale corresponds to 20° of arc distance.

of tests. First, global comparison of the observed and predicted SS-S residuals shows that ZT underpredicts observations by a factor of 2-3. Second, the slope b of the predicted SS and SS-S anomalies ($\delta t = a + b\tau^{1/2}$, where τ is the sea-floor age^{21,22}) is for ZT about 5 times smaller than for the observed residuals. The details of this experiment are shown in Fig. 2. Third, comparison of ~ 900 observed mantle-wave seismograms with synthetics computed for ZT shows that it provides only a minor improvement over a spherically symmetric Earth model, PREM (ref. 23). We use centroid moment tensor solutions for 21 earthquakes with magnitude $M_w \geq 6.5$ that occurred between July and December 1991^{24,25}. The median residual variance for PREM is 45%. Correction for lateral heterogeneity using ZT brings it down to 39%; the value is 36% for MDLSH (ref. 9), 31% for M84C and 30% for a model described below.

The dotted line in Fig. 1*d* corresponds to the predictions of a recent model SH8/WM13 (ref. 4; WM for short) obtained by inverting simultaneously waveform data⁵ (6,000 mantle-wave seismograms, 9,000 long-period body-wave seismograms) and differential travel-time data (5,600 SS-S (ref. 6) and 2,600 ScS-S (ref. 7) residuals). WM is a whole-mantle shear velocity model, in which radial variations are parameterized using Chebyshev polynomials up to degree 13 and horizontal variations by spherical harmonics up to degree 8. There is no explicit discontinuity between the upper and lower mantle.

The data set used to obtain WM is larger and more diverse than that used to obtain M84C (ref. 8) (2,000 mantle-wave seismograms), yet the main features of both models, as well as those of MDLSH (ref. 9), are similar. The same is true with respect to the lower-mantle structure^{3,9,13,19}. This means that modelling of the velocity anomalies in the mantle can be robust. Model ZT was derived using surface-wave data extending to periods shorter (75 s) than those used in derivation of M84C or

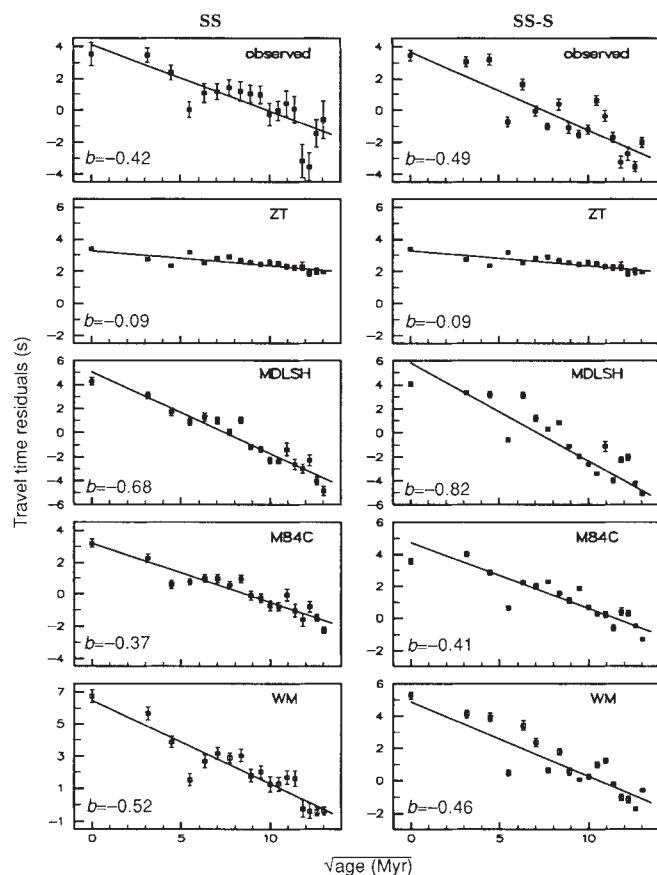


FIG. 2 Cooling of the oceanic plate with time leads to an increase in the seismic velocities and, consequently, variation in travel-time residuals as a function of the sea-floor age²². Comparison of the observed and predicted residuals is a good test of a globally integrated properties of the oceanic structure of a three-dimensional model. The SS residuals (left) and SS-S residuals (right) are shown as a function of the square root of the age of oceanic plate. From top to bottom: observations^{3,6}; predictions for ZT; predictions for MDLSH (ref. 9); predictions for M84C (ref. 8); predictions for WM (ref. 4). The slope predicted by model ZT is significantly shallower than that observed or predicted by other models.

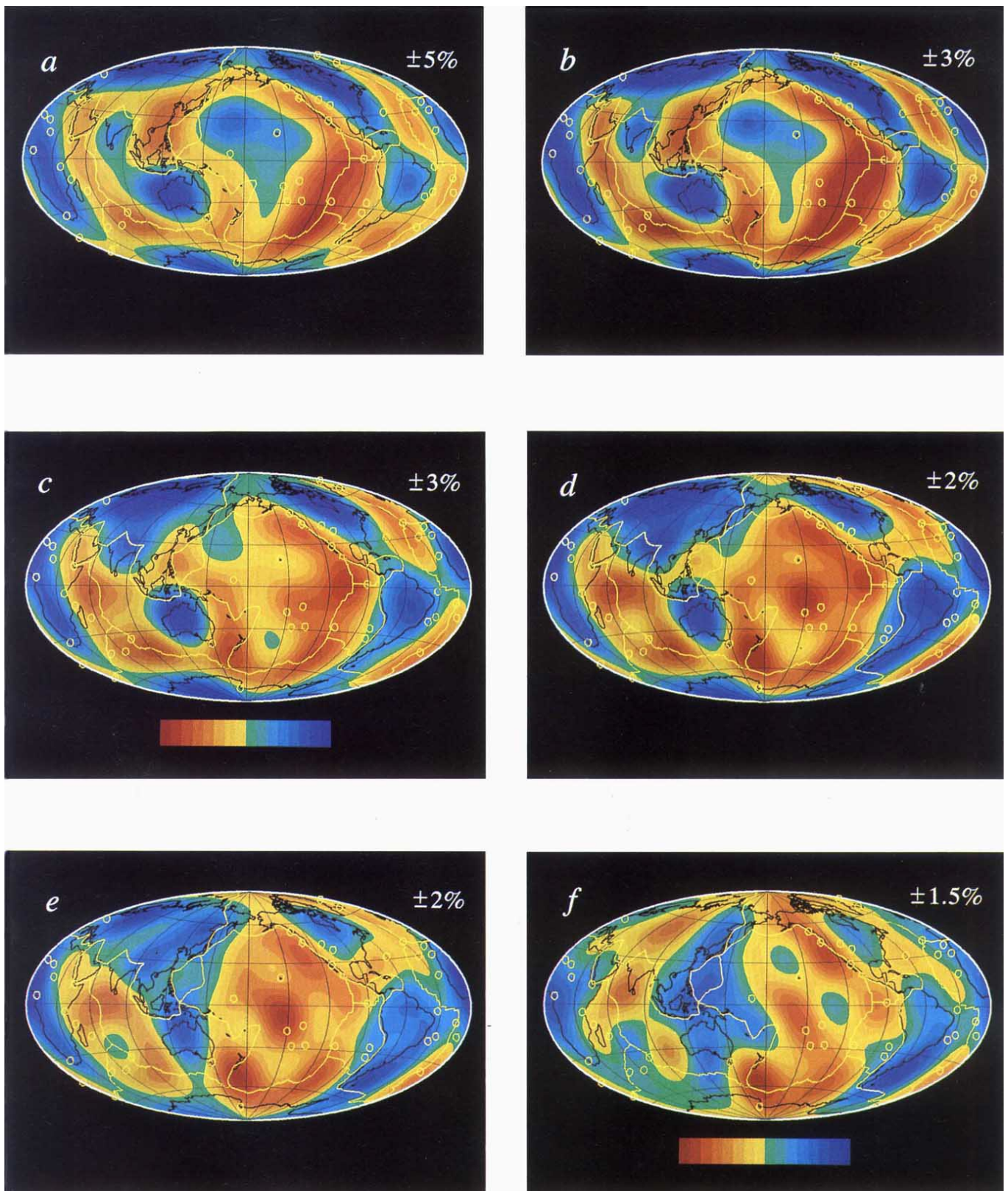


FIG. 3 Maps of velocity anomalies in model SH8/WM13 (ref. 4) at a depth of *a*, 50, (*b*) 100, (*c*) 200, (*d*) 300, (*e*) 400 and (*f*) 600 km. The colour code is shown at the bottom and the scale range for each map is indicated in the upper right-hand corner. Plate boundaries are shown in yellow and

hotspots are shown as circles. Except for the difference caused by the high-wavenumber features, these can be directly compared to Fig. 1*a-e* in ref. 11.

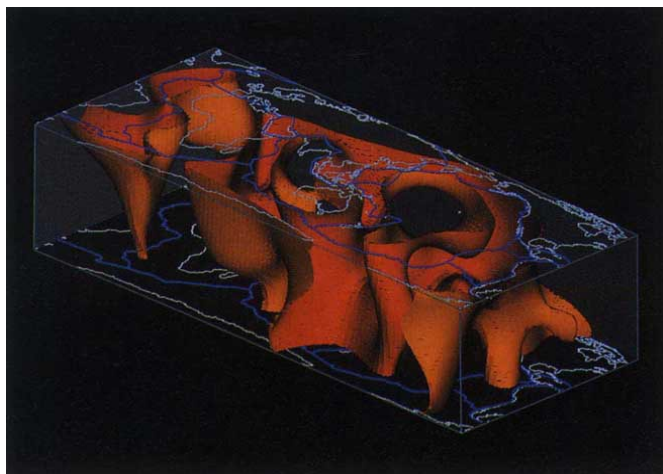


FIG. 4 The rectangular box spans 360° in longitude, split at 60° W, extends from 75° S to 75° N in latitude, and from the Moho to 660 km in depth; the vertical dimension is exaggerated 15 times. The box contains surfaces that correspond to -0.5% shear velocity anomaly in model WM; the inner surface contains even slower material and is shown in red; the outer surface is adjacent to material faster than -0.5% and is shown in amber. Plate boundaries are in blue. The figure shows a view from 135° azimuth (SE) from a point several thousand kilometres above the Earth's surface. The troughs associated with the mid-ocean ridges and back-arc basins are clearly seen. The anomalies associated with back-arc basins are shallow. Some of the features, such as the Pacific–Antarctic Ridge and parts of the Southwest Indian Ridge, extend from the surface to the 660-km discontinuity. Except for a feature in the South-Central Pacific, most of the anomalies extending below this depth are associated with the ridge structures at the surface.

WM (135 s). This, however, cannot be the cause of the discrepancy, as the short-period data are most sensitive to shallow structure, and this is the region where ZT and WM are most similar.

Figure 3 examines the depth expression of mid-oceanic ridges in the upper mantle of model WM. (Note that the scale for the colour code changes with depth, as the anomalies decrease in amplitude.) At a depth of 50 km all ridge structures are slower than average, with the EPR anomaly being the slowest (-5%); the ridge axes follow very closely the minima of the velocity field. The peak velocity perturbations decrease to about -4% at 100 km and -3% at a depth of 200 km. Some velocity lows associated with the back-arc basins of the western Pacific disappear between 100 and 150 km depth, as do the high velocities of the old (cold) oceanic lithosphere. This demonstrates that the depth extent of these shallow features is resolved by the model.

The ridges continue to be well defined at 200 km depth, with the velocity anomalies being larger for higher spreading rates; in ZT the dependence on spreading rate disappears by 100 km depth. The velocity anomalies associated with ridges remain continuous at 300 km depth, but the largest off-ridge anomalies become roughly equal in magnitude. At 400 km the largest negative anomalies are in the central Pacific and south of New Zealand, yet the ridges remain, on average, slower than normal, with the anomalies under the North and South Atlantic parts of the mid-Atlantic Ridge, Pacific–Antarctic Ridge, Southwest Indian Ridge and Carlsberg Ridge exceeding -0.5% and sometimes -1% . The situation is similar at 600 km, except that the velocity lows are no longer symmetric with respect to the ridge axes. A velocity low off the west coast of North America, which is part of the overall low-velocity trough near the surface,

becomes a distinct minimum at about 300 km and develops into one of the largest slow features at 600 km.

The depth extent of the positive velocity anomalies (Fig. 3) should also be noted; the fact that continents are uniformly faster than average at a depth as great as 400 km lends support to the hypothesis of deep 'continental roots'^{26,27}.

We have shown so far that the depth extent of the velocity anomalies associated with the mid-ocean ridges is significantly greater than the 100 km indicated by ZT. Ridge anomalies are distinct and continuous down to 300 km depth, and some features persist as deep as 600 km. The question occurs whether these anomalies are related to the pattern of lateral heterogeneity in the lower mantle. Certain features of recent whole-mantle models^{4,16} indicate that this may be the case. A three-dimensional view of a surface representing -0.5% shear velocity anomaly in the upper mantle is shown in Fig. 4: this surface intersects the 660-km discontinuity in several places.

Because of the more complicated three-dimensional geometry, the image of velocity anomalies in the lower mantle is not amenable to a simple test such as presented in Fig. 1. Also, the geometrical pattern is more complex than the relation between the circum-Pacific subduction zones and the location of the high velocity anomalies in the lower mantle¹⁴. Yet the hypothesis that some of the mid-ocean-ridge anomalies have a deep origin could be tested by carrying out seismic experiments using large portable arrays. Appropriate numerical simulations of mantle convection are also needed. Some answers may come from different branches of Earth sciences: for example, through studying the large-scale patterns of isotopic signatures in oceanic basalts. The similarity between the pattern of the Dupal anomaly^{28,29} and the pattern of low velocities in the lower mantle has been noted³⁰. □

Received 6 April; accepted 4 September 1992.

- Zhang, Y.-S. & Tanimoto, T. *Nature* **355**, 45–49 (1992).
- Zhang, Y.-S. thesis, California Institute of Technology (1991).
- Su, W.-J. & Dziewonski, A. M. *Nature* **352**, 121–126 (1991).
- Woodward, R. L., Forte, A. M., Su, W.-J. & Dziewonski, A. M. in *Proc. Union Symp. 12 IUGG* (eds Takahashi, E. & Jeanloz, R.) (Am. Geophys. Un., Washington DC, in the press).
- Woodhouse, J. H. & Dziewonski, A. M. *Phil. Trans. R. Soc. A* **329**, 291–308 (1989).
- Woodward, R. L. & Masters, G. *J. geophys. Res.* **96**, 6351–6377 (1991).
- Woodward, R. L. & Masters, G. *Nature* **352**, 231–233 (1991).
- Woodhouse, J. H. & Dziewonski, A. M. *J. geophys. Res.* **89**, 5953–5986 (1984).
- Tanimoto, T. *Geophys. J. int.* **100**, 327–336 (1990).
- Masters, G., Jordan, T. H., Silver, P. G. & Gilbert, F. *Nature* **298**, 609–613 (1982).
- Anderson, D. L., Tanimoto, T. & Zhang, Y.-S. *Science* **256**, 1645–1651 (1992).
- Tanimoto, T. & Zhang, Y.-S. *Geophys. Res. Lett.* **19**, 1567–1570 (1992).
- Dziewonski, A. M., Hager, B. H. & O'Connell, R. J. *J. geophys. Res.* **82**, 239–255 (1977).
- Dziewonski, A. M. *J. geophys. Res.* **89**, 5929–5952 (1984).
- Giardini, D., Li, X.-D. & Woodhouse, J. H. *Nature* **325**, 405–411 (1987).
- Dziewonski, A. M. & Woodward, R. L. in *Acoustical Imaging* (eds Emert, H. & Harjes, H.-P.) 785–797 (Plenum, New York, 1992).

- Hager, B. H., Clayton, R. W., Richards, M. A., Comer, R. P. & Dziewonski, A. M. *Nature* **313**, 541–545 (1985).
- Forte, A. M. & Peltier, W. R. *J. geophys. Res.* **92**, 3645–3679 (1987).
- Forte, A. M. & Peltier, W. R. *Geophys. Res. Lett.* **16**, 621–624 (1989).
- Forte, A. M., Dziewonski, A. M. & Woodward, R. L. in *Dynamics of the Earth's Deep Interior and Earth Rotation* (ed. Le Mouél, J.-L.) (Am. Geophys. Un., Geodyn. Ser., in the press).
- Parsons, B. & Slater, J. G. *Geophys. J. R. astr. Soc.* **82**, 803–827 (1977).
- Duschenes, J. D. & Solomon, S. C. *J. geophys. Res.* **82**, 1985–2000 (1977).
- Dziewonski, A. M. & Anderson, D. L. *Phys. Earth planet. Inter.* **25**, 297–356 (1981).
- Dziewonski, A. M., Ekström, G. & Saigani, M. *Phys. Earth planet. Inter.* **72**, 1–11 (1992).
- Dziewonski, A. M., Ekström, G. & Saigani, M. *Phys. Earth planet. Inter.* (in the press).
- Jordan, T. H. *Rev. Geophys. Space Phys.* **13**, 1–12 (1975).
- Jordan, T. H. *Nature* **274**, 544–548 (1978).
- Dupré, B. & Allegre, C. J. *Nature* **303**, 142–146 (1983).
- Hart, S. R. *Nature* **309**, 753–757 (1984).
- Castillo, P. R. *Nature* **336**, 667–670 (1988).

ACKNOWLEDGEMENTS. We thank Y.-S. Zhang for providing the numerical details of his model. This research was supported by the NSF.

Stable isotope evidence for entry of sewage-derived organic material into a deep-sea food web

Cindy Lee Van Dover*, J. F. Grassle†, Brian Fry‡, Robert H. Garritt§ & Victoria R. Starczak§

* Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

† Institute for Marine and Coastal Sciences, Cook College, Rutgers University, New Brunswick, New Jersey 08903, USA

‡ Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

§ Applied Ocean Physics and Engineering Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

CHRONIC pollution of the open ocean has occurred since 1986 through disposal of municipal sewage sludge at a deep-water (~2,500 m) dumpsite off the coast of New Jersey. Dispersal and dilution of sewage particulates in surface waters were presumed to be sufficient to prevent or minimize accumulation of detectable amounts of sewage-derived material on the sea floor. Using stable isotope ratios of carbon, nitrogen and sulphur as tracers of sewage-derived organic material, we show here that this material reaches the sea floor and enters the benthic food web, specifically through surface-deposit feeding activities of the urchin, *Echinus affinis* and the sea cucumber, *Benthodytes sanguinolenta*.

Megascale pollution of the open ocean 185 km off the coast of New Jersey was maintained from March 1986 to July 1992 through dumping of municipal sewage sludge at a rate of 8–9 million wet metric tons per annum¹. The potential for perturbation and/or environmental degradation of benthopelagic and benthic ecosystems within and downstream of the dumpsite, designated Deep-Water Dumpsite 106 (DWD-106), was originally de-emphasized by planners, whose initial monitoring

TABLE 1 Isotope composition of sewage sludge and of open-ocean, phytoplankton-derived particulate organic material

$\delta^{13}\text{C}$	$\delta^{15}\text{N}$	$\delta^{34}\text{S}$	Location	Reference
Sewage-derived organic material				
-24.7	-1.1	+2.3	Middlesex, NJ	this study
-23.2	+6.1	+2.3	Bergen, NJ	this study
-21.4	+7.2	+3.5	Yonkers, NY	this study
-23.0	+3.2	+4.0	NY/NJ	24
-26.0			NY Bight	12
	+2.5	0.0	Whites Point, CA	25
-16.5	+1.8		Whites Point, CA	15
-23.5			Whites Point, CA	26
-23.7		+3.0	Providence, RI	this study
Phytoplankton-derived particulate organic material				
-21.7	+6.1		N. Atlantic 38° 57.79' N 71° 56.7' E sediment top 1 cm 2,419 m	this study
-21.6			Sedimentary POC Western N. Atlantic 3,471 m	27
+5 to +15			Deepwater Pacific PON	28
+5 to +7			Sargasso Sea Nisken and Pump Sample PON 1,000 m	29
		+17 to +21	Marine algae POS	16, 23, 30

strategies focused exclusively on the water column^{2,3}. But subsequent characterization of sludge particle sizes and settling dynamics⁴ suggests that initial interpretations of a benthically innocuous fate for sewage sludge may be erroneous. New models using historic current meter and particle settling-velocity data⁵ predict measurable sea-bed loading of sewage-derived particulates up to 350 km downstream of the dumpsite. Recent studies of sediments in the region indicate that levels of sewage indicators are increased in the dumpsite area^{6–9}.

The magnitude of organic loading 50 km downstream of the dumpsite is estimated to be of the same order as seasonal pulses of phytodetritus reaching the abyssal sea floor after surface blooms (F. Sayles, personal communication, based on estimates of sewage flux to the sea floor^{3,5} and background flux of organic carbon¹⁰). Such a doubling of food resources at DWD-106 is likely to have a detectable effect on the deep-sea biotic community. Evidence for entry of sewage-derived organic material into the deep-sea food web would provide a conclusive test of this hypothesis.

We chose to exploit the natural stable isotope compositions of organic carbon, nitrogen and sulphur in sewage sludge as tracers for sewage-derived organic matter (SDOM)^{11–15}. If sewage sludge does not enter the abyssal food web, megafaunal invertebrate and fish species trawled from areas within and immediately downstream of the disposal site should have similar carbon, nitrogen and sulphur isotope compositions as organisms from reference areas upstream of the site along similar depth contours and sediment types. This null hypothesis assumes that the only difference between dumpsite and reference areas is the potential for sewage-derived organic enrichment within and downstream of the dumpsite, an assumption consistent with submersible observations and our present understanding of the seafloor environment within the region. To establish end-member values (Table 1), we analysed carbon and nitrogen isotope compositions of plankton-derived organic material (PDOM) in the surface layer of sediment from a site upstream

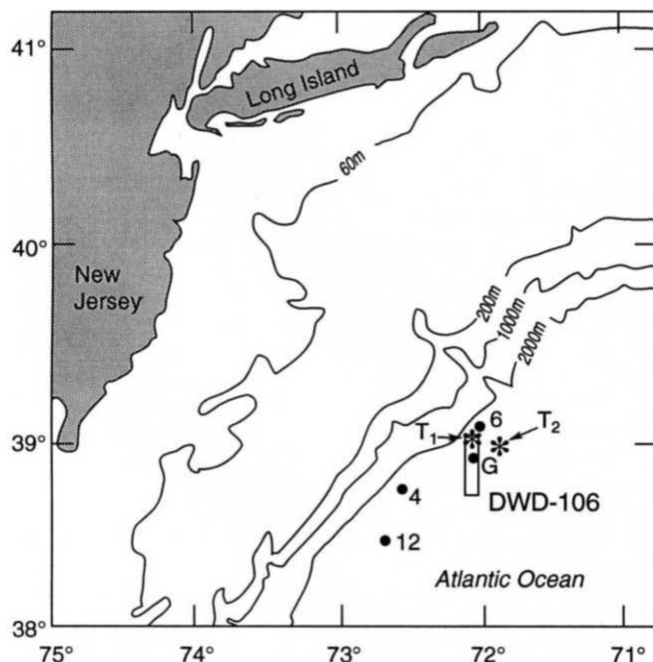


FIG. 1 Location of Deep-Water Dumpsite 106 off the coast of New Jersey. Asterisks indicate locations of 1989 trawls. 1990 trawl locations are designated by station (6, G, 4, 12). One trawl per site was taken each year.

TABLE 2 Carbon, nitrogen and sulphur isotope compositions (‰) of megafaunal species collected from dumpsite and reference trawls

Species		Tissue	Site	$\delta^{13}\text{C}$	$\delta^{15}\text{N}$	$\delta^{34}\text{S}$
<i>Flabellum angulare</i>	Solitary coral	Soft parts	DWD-106	-19.9±0.3 (5)	+10.3±0.5 (5)	+17.6±0.3 (3)
			Reference	-19.7±0.4 (8) [0.42]	+10.5±0.3 (8) [0.33]	+18.1±1.6 (3) [0.62]
<i>Echinus affinus</i>	Sea urchin	Gonad	DWD-106	-22.1±1.1 (10)	+5.5±1.2 (10)	+15.3±1.2 (9)
			Reference	-21.0±0.6 (3) [0.13]	+9.9±0.3 (3) [<0.000]**	+17.4±0.6 (3) [0.01]*
<i>Brissopsis mediterranea</i>	Heart urchin	Gonad	DWD-106	-21.7 (1)	+11.0 (1)	—
			Reference	-21.1 (1)	+10.9 (1)	—
<i>Pectinaster filholi</i>	Seastar	Gonad	DWD-106	-20.2±0.6 (3)	+11.8±2.0 (3)	+16.9±0.3 (3)†
			Reference	-21.0±0.5 (3) [0.11]	+12.6±3.1 (3) [0.74]	+15.7±1.6 (3)† [0.34]
<i>Benthodytes sanguinolenta</i>	Sea cucumber	Body wall	DW-106	-17.1±0.8 (3)	+7.7±0.8 (3)	+17.1±0.4 (3)
			Reference	-14.7±0.7 (3) [0.02]*	+11.2±0.5 (3) [0.003]**	+18.7±0.2 (3) [0.002]**
<i>Hedingia albicans</i>	Sea cucumber	Body wall	DWD-106	—	—	+17.6±0.3 (3)
			Reference	—	—	+17.9±0.3 (3) [0.38]
<i>Glyphocrangon sculpta</i>	Shrimp	Muscle	DWD-106	-16.0 (1)	+13.0 (1)	—
			Reference	-15.9 (1)	+13.2 (1)	—
<i>Munidopsis</i> sp	Squat lobster	Muscle	DWD-106	-17.7 (1)	+10.5 (1)	—
			Reference	-17.6 (1)	+10.8 (1)	—
<i>Antimora rostrata</i>	Deep-sea cod	Muscle	DWD-106	-18.0 (1)	+12.4 (1)	—
			Reference	-18.5±0.8 (3)	+12.1±0.3 (3)	—
<i>Lionurus carpinus</i>	Rat tail	Muscle	DWD-106	-18.0±0.6 (2)	+12.8±0.7 (2)	—
			Reference	-17.5 (1)	+12.7 (1)	—

Values are given as mean $\pm$ s.d. (n); analytical error is $\pm 0.1\%$ except for small sulphur samples indicated by †, where the analytical error is $\pm 0.25\%$. When $n \geq 3$, isotope compositions were compared between dump site and reference site using a two sample *t*-test. Significant results are indicated with asterisks; *P* values are given in brackets. Statistical analysis assumed that individuals were sampled randomly from each site.

of the dumpsite; sulphur and additional carbon and nitrogen isotope data for PDOM are reported from the literature. Isotope composition of sewage sludge was analysed in February 1991 samples from contributing metropolitan New York and New Jersey treatment facilities (Table 1). Comparison of PDOM and SDOM indicates that sulphur should provide the strongest signal ($|\Delta\delta_{\text{PDOM-SDOM}}| = 11.5$ to 18.7%), carbon the weakest signal ($|\Delta\delta_{\text{PDOM-SDOM}}| = 0.3$ to 3.1%), with nitrogen intermediate ($|\Delta\delta_{\text{PDOM-SDOM}}| = 1.1$ to 7.2%). The large difference in the sulphur isotope composition of PDOM and SDOM is attributable to the +2 to +6‰ $\delta^{34}\text{S}$ composition of continental vegetation and the +17 to +21‰ values of marine plankton¹⁶.

We initially screened ten megafaunal species common to trawls from DWD-106 and a reference site collected in September 1989 (Fig. 1) for carbon, nitrogen, and/or sulphur isotope composition (Table 2). For five species, isotope analyses were done on tissues from 3 or more individuals per site. Carbon isotope compositions of dumpsite and reference pairs for four of these species were not significantly different. There was a significantly more negative $\delta^{13}\text{C}$ composition in the sea cucumber, *B. sanguinolenta*. The $\sim 2.5\%$ depletion of ^{13}C in the dumpsite *B. sanguinolenta* is consistent with a contribution of SDOM to its diet. Between-species differences in $\delta^{13}\text{C}$ values may be attributed to the biochemical composition of the tissues analysed: lipid-rich tissues such as gonad and the soft parts of the solitary coral (predominantly gonad) are isotopically lighter ($\delta^{13}\text{C} = -19.7$ to -21.9%) than protein-rich tissues of body wall and muscle ($\delta^{13}\text{C} = -14.7$ to -18.5%) owing to discrimination towards ^{12}C during lipogenesis^{17,18}.

Eight of ten species analysed for nitrogen and/or sulphur isotope composition show no difference between dumpsite and reference individuals within species (Table 2). In two species, *E. affinus* and *B. sanguinolenta*, there are significant differences in $\delta^{15}\text{N}$ and $\delta^{34}\text{S}$ values between dumpsite and reference individuals. These differences are consistent with contributions of ^{15}N - and ^{34}S -depleted SDOM in the diet of dumpsite individuals. Both *E. affinus* and *B. sanguinolenta* are surface-deposit feeders. In contrast, the other megafaunal species surveyed are variously suspension feeders (*Flabellum angulare*), subsurface deposit feeders (*Brissopsis mediterranea*, *Hedingia*

albicans), or motile predators or scavengers (*Glyphocrangon sculpta*, *Munidopsis* sp., *Antimora rostrata*, *Lionurus carpinus*, and possibly *Pectinaster filholi*). Suspension feeders feed in the water column where SDOM does not accumulate. In subsurface sediments, concentrations of SDOM are reduced to undetectable levels, based on measurement of sewage indicators^{6,7}. Predatory or scavenging shrimp, crabs and fish feed on a variety of benthic invertebrates, some of which may have access to SDOM. The foraging ambit of these motile predators, however, may take them beyond the area influenced by sludge disposal, thereby reducing their opportunities for uptake of SDOM. Surface deposit-feeders, *E. affinus* in particular, are known to be attracted to aggregations of food¹⁹, including seasonal pulses of organic matter^{20,21} and, as shown here, SDOM should be most readily available to surface-deposit-feeding species.

Individuals of *E. affinus* from the reference area, where only organic material from natural sources is available, show less variability in isotope composition than dumpsite urchins (Table 2). Isotope variability within the dumpsite population may reflect microgeographic differences in deposition and availability of SDOM versus PDOM⁶⁻⁹. There may also be an age-dependent variation if turnover rate of tissues within an individual is longer than the period of dumping; tissues of individuals populating the dumpsite area before the start of dumping may reflect a mixture of pre- and post-dumping trophic conditions. We examined this latter possibility by analysing $\delta^{34}\text{S}$ of urchin size series trawled in August 1990 from dumpsite and reference areas. We expect and find no linear relationship between size and $\delta^{34}\text{S}$ composition in reference urchins and a significant positive linear relationship between size and $\delta^{34}\text{S}$ composition of dumpsite urchins, reflecting an increased contribution of SDOM to smaller, more rapidly growing individuals (Fig. 2a).

Using a simple, two-source mixing model, we can estimate the relative importance of sewage-derived sulphur compounds in the metabolism of dumpsite urchins: $\delta^{34}\text{S}_{\text{dumpsite urchin}} = f\delta^{34}\text{S}_{\text{sewage sludge}} + (1-f)\delta^{34}\text{S}_{\text{phytodetritus}}$, where *f* is the proportion of sewage-derived organic sulphur in the diet of the urchin and $\delta^{34}\text{S}$ values for urchins, sludge and phytodetritus are 12.2, 3.0 and 17.4‰, respectively. Because we see little difference between

$\delta^{34}\text{S}$ of phytoplankton and of reference urchins consuming PDOM, we infer that there is little fractionation of sulphur isotopes in organic material as it sinks through the water column. Other laboratory and field studies indicate that little sulphur isotope fractionation occurs during decomposition of organic material or in off-shore ecosystems^{22,23}. Further, $\delta^{34}\text{S}$ of sediment recovered from traps moored 250 m above the bottom

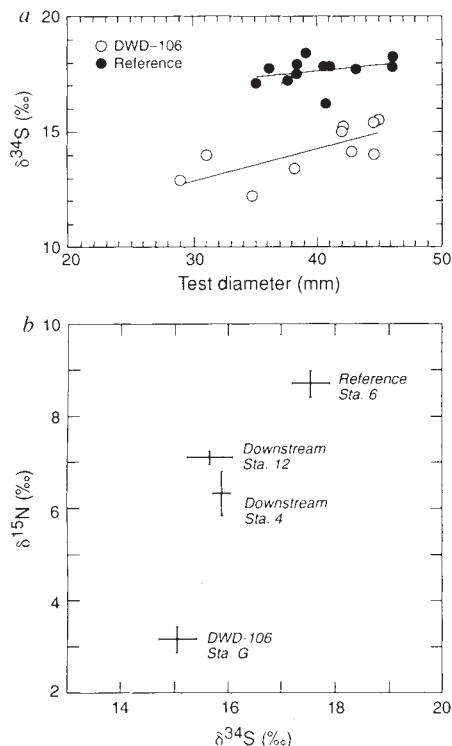


FIG. 2 a, $\delta^{34}\text{S}$ (‰) of urchin (*Echinus affinis*) gonads versus urchin test diameter (mm) for 1990 trawl samples. Reference urchins had test diameters ranging from 35 to 46 mm, corresponding to an age range of ~15 to 25 years, based on a growth model generated for a population of *E. affinis* in the Rockall Trough (northeastern North Atlantic) at 2,200 m³¹. Under the same growth model, dumpsite urchins ranging in diameter from 29 to 45 mm correspond to an age range of 10 to 25 years. Thus all of the urchins analysed are likely to have experienced pre-dumping trophic conditions, with smallest animals experiencing a greater proportion of growth since the initiation of dumping in 1986. Isotope compositions are expressed using the delta notation¹⁶ and were determined using standard techniques^{32,33}. Reference regression equation and statistics: $y = 15.647 + 0.049x$; $F_{1,10} = 0.97$; $R^2 = 0.088$; $P = 0.3488$. Dumpsite regression equation and statistics: $y = 8.807 + 0.136x$; $F_{1,8} = 8.47$; $R^2 = 0.514$; $P = 0.02$. A greater number of smaller individuals were analysed from the dumpsite than from the reference site. A regression analysis was performed using only individuals with test diameters >34 mm. The resulting regression ($y = 2.973 + 0.272x$; $R^2 = 0.712$; $F_{1,6} = 14.81$; $P = 0.008$) was significant, indicating that the results in the original analysis are robust to the deletion of individuals 34 mm or smaller. b, $\delta^{34}\text{S}$ vs $\delta^{15}\text{N}$ (‰, $\pm$ s.e.) for urchin (*E. affinis*) gonads collected in 1990 trawls; $n = 5$ for all data sets except Station 12, where $n = 4$ for sulphur isotope analysis. Station locations are shown in Fig. 1. Mean test diameters ($\pm$ s.d.): Station G: 43.8 ± 1.3 mm; Station 6: 43.3 ± 2.7 mm; Station 4: 36.4 ± 2.7 mm; Station 12: 44.0 ± 2.5 mm. Tukey's Multiple Comparisons test on the mean $\delta^{34}\text{S}$ and $\delta^{15}\text{N}$ values from the four stations sampled shows that $\delta^{34}\text{S}$ and $\delta^{15}\text{N}$ values of dumpsite and reference urchins are significantly different and that urchins from downstream stations 4 and 12 had nitrogen isotope compositions intermediate to those of dumpsite and reference urchins. These results are consistent with modelling predictions of significant but decreasing flux of SDOM to the seafloor at distances to the southwest of 50 km and more⁵. There was no difference in mean $\delta^{34}\text{S}$ among urchins sampled from the dumpsite and from intermediate stations 4 and 12. Mean $\delta^{15}\text{N}$ values of dumpsite urchins decreased significantly from 5.5‰ in 1990 (2-sample t -test; $t = 5.04$; $P = 0.0002$; d.f. = 13) but no significant difference in $\delta^{34}\text{S}$ was detected in the one-year interval ($t = 0.33$; $P = 0.750$; d.f. = 11).

within 20 nmi of the dumpsite in 1991 approached $\delta^{34}\text{S}$ values of the sludge dumped at the site²⁴. We thus use the mean $\delta^{34}\text{S}$ for sewage sludge from New York and New Jersey treatment plants (Table 1) as one end-member source in the mixing model. The second end-member, $\delta^{34}\text{S}_{\text{phytodetritus}}$, is approximated by the average $\delta^{34}\text{S}$ of reference urchins in 1989 (Table 2). Using this model, we find that as much as 35% of the sulphur in gonadal tissues of dumpsite urchins may be derived from SDOM. This is likely to be a conservative value, inasmuch as the most recent model of sludge particle flux⁵ shows a maximum flux in a region 10 km west of the dumpsite, a location from which we have not yet collected urchins.

In August 1990, additional *E. affinis* were collected by trawls from a reference site (station 6) north and upstream of the dumpsite, from a location in the dumpsite (station G), and from two stations (stations 4, 12) located about 70 km to the southwest, or downstream of the dumpsite (Fig. 1). From each of these stations, we selected urchins of roughly equal size and analysed the $\delta^{15}\text{N}$ and $\delta^{34}\text{S}$ compositions of their gonads. We again found clear, two-dimensional separation of dumpsite and reference urchins (Fig. 2b). Within-site variability was considerably reduced in the 1990 samples, which most likely reflects the selection of more uniform-sized urchins for analysis.

Based on these isotope data, we conclude that SDOM reaches the deep-sea floor and enters the benthic food web as a result of consumption by surface-deposit feeders, specifically the sea cucumber *B. sanguinolenta* and the urchin *E. affinis*. Our use of stable isotope tracers was limited to large, readily collected megafaunal species with sufficient organic tissue to allow analysis on an individual basis. We anticipate that other surface-deposit-feeding organisms, including infaunal macroinvertebrates (such as polychaetes, molluscs, crustaceans) will prove to have access to SDOM.

Contrary to earlier assumptions, chronic disposal of sewage sludge in the open ocean does impact the benthic ecosystem underlying and downstream of the disposal site. It is likely that long-term disposal programs in the open ocean will result in a restructuring of the benthic community in favour of species such as *E. affinis* and *B. sanguinolenta*, which can readily exploit particulate organic enrichments. With the cessation of dumping in July 1992, the data reported here serve as a critical benchmark for assessment of recovery of the benthic food web to pre-dumping conditions. □

Received 11 August; accepted 22 September 1992.

- Robertson, A. & Redford, D. *Eos Trans. Am. Geophys. Union* **73**, 164 (1992).
- O'Connor, T. P., Walker, H. A., Paul, J. F. & Bierman, V. J. *Mar. env. Res.* **16**, 127–150 (1985).
- O'Connor, T. P., Okubo, A., Champ, A. & Park, K. P. *Can. J. Fish. Aqu. Sci.* **40** (suppl. 2), 228–241 (1985).
- Lavelle, J. W., Ozturgut, E., Baker, E. T., Tennant, D. A. & Walker, S. L. *Envir. Sci. Technol.* **22**, 1202–1207 (1988).
- Fry, V. A. & Butman, B. *Envir. Res.* **31**, 145–160 (1991).
- Bothner, M. H., d'Angelo, W. M. & Grassie, J. F. *Eos Trans. Am. Geophys. Union* **72**, 83 (1992).
- Takada, H., Tripp, B. W., Clifford, C. H., Farrington, J. & Bothner, M. *Eos Trans. Am. Geophys. Union* **72**, 83 (1992).
- Hill, R. T., Knight, I., Anikis, M., Straube, W. L. & Colwell, R. R. *Eos Trans. Am. Geophys. Union* **72**, 83 (1992).
- Grassie, J. F. *Eos Trans. Am. Geophys. Union* **72**, 84 (1992).
- Biscaye, P. E., Anderson, R. F. & Deck, B. L. *Continental Shelf Res.* **8**, 855–904 (1988).
- Sweeney, R. E. & Kaplan, R. R. *Mar. env. Res.* **3**, 215–224 (1980).
- Burnett, W. C. & Schaeffer, O. A. *Est. Coastal Mar. Sci.* **11**, 605–611 (1980).
- Gearing, P. J., Gearing, J. N., Maughan, J. T. & Oviatt, G. A. *Envir. Sci. Technol.* **25**, 295–301 (1991).
- Rau, G. H., Sweeney, R. E. & Kaplan, I. R. *Est. Coastal Shelf Sci.* **13**, 701–707 (1981).
- Spies, R. B., Kruger, H., Ireland, R. & Rice, D. W. Jr. *Mar. Ecol. Prog. Ser.* **54**, 157–170 (1989).
- Peterson, B. J. & Fry, B. A. *Rev. Ecol. Systematics* **18**, 293–320 (1987).
- Parker, P. L. *Geochim. cosmochim. Acta* **28**, 1155–1164 (1964).
- van der Merwe, N. J. *Am. Sci.* **70**, 596–606 (1982).
- Grassie, J. F., Sanders, H. L., Hessler, R. R., Rowe, G. T. & McLellan, T. *Deep Sea Res.* **22**, 457–481 (1975).
- Tyler, P. A. *Oceanogr. Mar. Biol. A. Rev.* **26**, 227–258 (1988).
- Tyler, P. A. & Gage, J. D. *Deep Sea Res.* **31**, 387–402 (1984).
- Mekhtiyeva, V. L., Parkina, R. G. & Gavrilov, Y. Y. *Geochim. Int.* **13**, 82–87 (1976).
- Fry, B. *Limnol. Oceanogr.* **33**, 1182–1190 (1988).
- Hunt, C. D., Ginsburg, L., West, D. & Redford, D. *Eos Trans. Am. Geophys. Union* **73**, 165 (1992).
- Sweeney, R. E., Kalil, E. K. & Kaplan, I. R. *Mar. Envir. Res.* **3**, 225–243 (1980).
- Myers, E. P. thesis, Calif. Inst. Technol. (1974).
- Sayles, F. L. & Curry, W. B. *Geochim. cosmochim. Acta* **52**, 2963–2978 (1988).
- Saino, T. & Hattori, A. *Deep Sea Res.* **34**, 807–827 (1987).
- Altabet, M. A. *Deep Sea Res.* **35**, 535–554 (1988).

30. Kaplan, I. R., Emery, K. O. & Rittenberg, S. C. *Geochim. cosmochim. Acta* **27**, 297–331 (1963).
 31. Gage, J. D. & Tyler, P. A. *Mar. Biol.* **90**, 41–53 (1985).
 32. Fry, B. *et al. Deep Sea Res.* **38**, 51003–51019 (1991).
 33. White, J. R., Gubale, C. P., Fry, B., Owen, J. & Mitchell, M. J. *Geochim. cosmochim. Acta* **53**, 2547–2559 (1989).

ACKNOWLEDGEMENTS. We thank R. Petrecca, the National Marine Fisheries Service, and the scientists and crews of *RVs Atlantis II* and *Delaware* for help in collecting specimens; D. Pabst and colleagues for helping us obtain samples of sewage sludge; M. Bothner for sediment samples; D. Pawson and B. Hecker for assistance in identification of organisms; and K. Tholke for isotope analysis. Parts of this work were sponsored by the NOAA Sea Grant College Program Office, Department of Commerce, and WHOI Sea Grant; additional support was provided by NOAA-NURP (J.F.G.).

Sexual dimorphism and distorted sex ratios in spiders

Fritz Vollrath* & Geoff A. Parker†

* Department of Zoology, University of Oxford, Oxford OX1 3PS, UK;
 Zoologisches Institut, Rheinsprung 9, CH 4051 Basel, Switzerland;
 Smithsonian Tropical Research Institute, Box 20272, Balboa, Panama
 † Department of Environmental and Evolutionary Biology,
 University of Liverpool, Liverpool L69 3BX, UK

SEXUAL dimorphism in body size is widespread in the animal kingdom. Whereas male gigantism has been studied and explained extensively^{1,2}, male dwarfism has not. Yet it is neither rare^{3–7} nor without theoretical interest^{8,9}. Here we provide experimental and comparative data on spiders to support the theory that dwarf males are associated with high differential adult mortality, with males at much greater risk. Species with sedentary (low-risk) females have dwarf, roving (high-risk) males. Life-history theory could readily explain dwarfing if juvenile, but not adult, male mortality were large. We present a new model in which high mortality of searching mature males reduces the adult sex ratio (males: females), relaxing male–male competition and reducing the importance of male body size to favour dwarfing by early maturation. Early maturity also reduces male juvenile mortality and thus opposes adult mortality. This provides a mechanism that buffers skews in adult sex ratio and which is quite distinct from Fisher's principle¹⁰ and allied mechanisms^{9,11} for the primary sex ratio.

Spiders have long provided a classic example of sexual dimorphism in body size⁸, and provide an ideal opportunity for empirical study. They comprise a number of families with widely differing life histories and differing extents of sexual dimorphism, which can be highly variable even within a single species^{3,12}. Thus spiders enable us to analyse the ecology of sexual dimorphism (with empirical investigations of particular species or ecotypes), as well as study its evolution (by using the comparative method¹³ to juxtapose different species or ecotypes).

The golden orb web-building spider *Nephila clavipes* has males which are tiny in comparison to their females³. Detailed life-history studies^{14,15} of this highly dimorphic animal reveal that the reproductive success of large males is much higher than small males. Yet large males are in the minority among their peers, and are much smaller than the females¹². Unidentified selection mechanisms apparently favour small or even dwarfish size in *Nephila* males. It was suggested that female spiders, by cannibalizing courting males, have actively selected for small male size^{4,16}. We argue that selection for dwarfism results from reduced intrasexual competition through high differential adult mortality between the sexes. This differential is primarily associated with males searching for sedentary females, though it will be exacerbated by cannibalism.

At our field study site in Panama, we found evidence for high mortality of immature *Nephila* (Fig. 1a). Adult females experienced much lower mortality rates whereas adult males searching for females¹² had even higher mortality rates. The large difference in mortality of adult males and females has important implications for the adult sex ratio in *Nephila*. From

the mortality data shown in Figs 1 and 2, we conclude that males mature earlier and that the mature size is smaller in females, but that mortality of the roving male exceeds that of the sedentary female. Alternative explanations for the observed phenomenon were discounted: males were not localized in areas of high population density, nor was the mortality bias reversed in such areas. The data indicate that dimorphism in adult size may be associated with mortality bias differences brought about by dissimilar adult life styles.

Our comparative studies provide evidence for such an association. Correlation of male and female body size demonstrates that most spiders show sexual dimorphism (Fig. 3a), with males

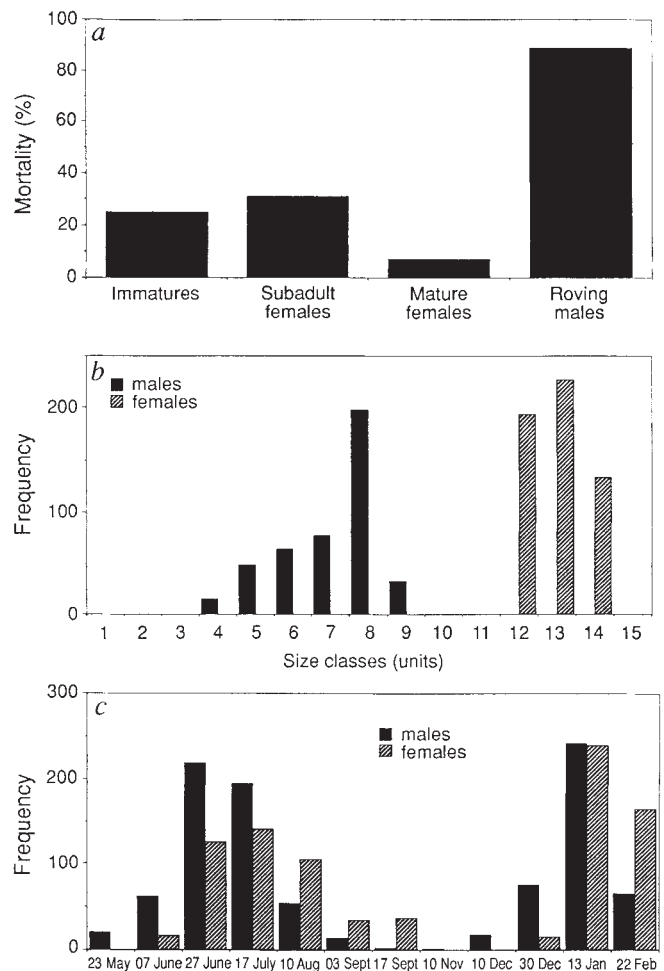


FIG. 1 Mortality, size distribution and maturation of *Nephila clavipes* in a rich habitat (Panama, Cerro Galera Hill Road). *a*, Mortality of immatures (sampled in 3 instars), females and males. Mortality of immatures ($n=791$), subadult females ($n=451$) and mature females ($n=355$) was calculated from a cohort study (June–September) where a given site was visited every 20 days. Data of male mortality ($n=138$) came from a capture–recapture study with marked individuals (daily visits to the site, after 10 days no more males were recaptured)¹². Note that male mortality is shown as % dying in 10 days whereas the other groups are shown as % dying in 20 days; hence the differences in mortality were even more pronounced than depicted. *b*, Size distribution of males and females classified in the field using a chart (1.5 units $\sim$ 1 mm). Size (patella–tibia measured under a binocular microscope) ranged in males from 2.3 to 8.1 mm with a mean ($\pm$ s.d.) of 4.9 ± 1.5 mm ($n=24$); dominant males were larger: 6.1 ± 0.9 mm ($n=10$). Males weighed 19.5 ± 15.6 mg. Female size ranged from 9.5 to 17 mm with a mean of 13 ± 2 mm, and at maturation they weighed 331.4 ± 184.8 mg ($n=10$). *c*, Timing of maturation of subadult and adult male and females in the field. In the laboratory, males had a significantly slower rate of growth, but still matured earlier than females. This was also highly significant ($P > 0.005$) and observed whether the diet was rich (males 33.3 ± 7.2 days, $n=11$; females 55.3 ± 4.6 days, $n=4$) or poor (males 82 ± 17.2 days, $n=12$; females 143 ± 5.8 days, $n=3$).

being typically smaller. The degree of dimorphism varies between the different arachnid families, being most pronounced in those families that typically forage as sit-and-wait predators, and less pronounced in active and roaming hunters. This difference is illustrated in Fig. 3b and c, using two pairs of families, each pair matched for female size but contrasting in adult lifestyle. The actively hunting wolf and jumping spiders show little dimorphism and similar adult life styles, whereas the sedentary orb web-building spiders and the similarly sessile (yet webless) crab spiders have males that are much smaller and much more mobile than their respective females. Indeed, we can find sedentary spiders like the webless, but sedentary, bolas spider, *Mastophora*¹⁷, or the unrelated tangle-web spider, *Tidarren*¹⁸, where males have undergone runaway evolution for dwarfism and mature inside the egg sac, clearly a derived character. The males of *Tidarren* are so small that they cannot even carry their full complement of mating implements and consequently amputate one of their two pedipalps. The Linyphiids are the main exception to the hypothesis. They are web spiders that have large males. But their mode of mating includes mate guarding with fierce male-male competition¹⁹, thus strong sexual selection seems to punish small size despite male-biased mortality²⁰.

The two spider life-history patterns are summarised in Fig. 4a. In type-A patterns, males and females have similar sizes and mortality rates throughout life. In type B males are smaller (they mature earlier than females) and have higher adult mortality. Current life-history theory would readily explain the small males in B if juvenile male mortality were higher than that of females,

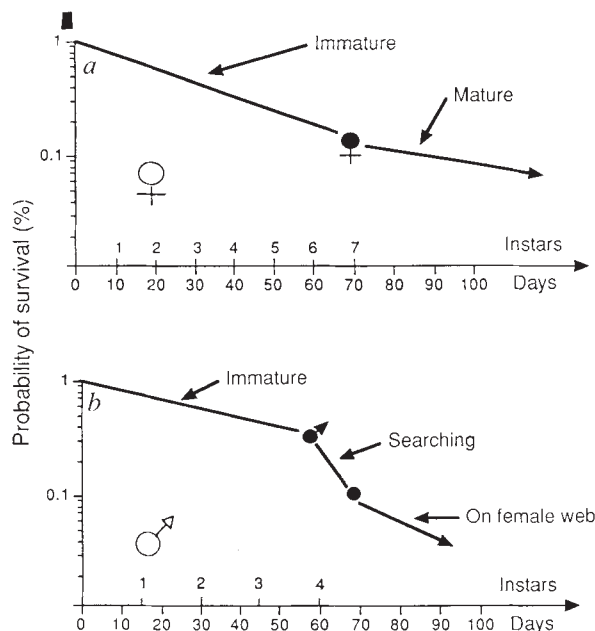


FIG. 2 Diagrammatic representation of mortality in a hypothetical spider with sexual dimorphism, based on data for *Nephila*. Gender symbols indicate the timing of maturation. *a*, Mature females have a lower mortality than immature females. *b*, Mature males have a higher mortality when searching for a female than on a female's web. The circle indicates the point on which the male finds a female. The primary sex ratio was unity: 142 spiderlings, selected at random from 2 egg sacs and raised without loss, yielded sex ratios (proportion male) of 0.49 and 0.47. The sex ratio of immatures in the field was 0.50 for instar 4 ($n=123$) and 0.48 for instar 5 ($n=93$). Once adult, males no longer build webs, and the adult sex ratio had to be estimated. At high population densities, when webs interconnect, male travel mortality was low and we observed an adult sex ratio (male:female) of 0.75 ($n=711$). Commonly, population density was much lower¹², and the males travelled considerable distances (upwards of 10 metres) with all the associated risks. In a typical *Nephila* habitat (secondary growth in recent clearings) we observed an operational sex ratio of 0.49 ($n=59$) during the height of the mating season. In another similar habitat, sampled 9 times over two spider generations, we observed an average sex ratio of 0.52 ($n=1,076$).

but adult mortality should not affect the timing of maturity or adult size. Each sex switches from growth to reproduction at an optimal age, t_m^* for males and t_f^* for females, depending on how size affects fitness for each sex. Classically²¹ it is held that the female optimum depends on how size increases fecundity, whereas the male optimum depends on how size increases mating success. Both of these size benefits are traded off against juvenile mortality; longer juvenile life increases size, but reduces chances of reaching adulthood. But it is not clear how adult mortality rates could affect the optimal timing of maturity. We suggest a new form of link, arising through the influence of operational sex ratio²² on the benefits of size to male fitness.

Our model analyses size dimorphism with differential adult mortality. Mortality is random, but dependent on lifestyle (for example mobility pattern)¹²; instantaneous death rates for juveniles are d_{jm} , d_{jf} and for adults are d_{am} , d_{af} (for males and females, respectively). For simplicity, growth is linear¹⁵. We seek the optimal age (and hence size) at which fitness (survival

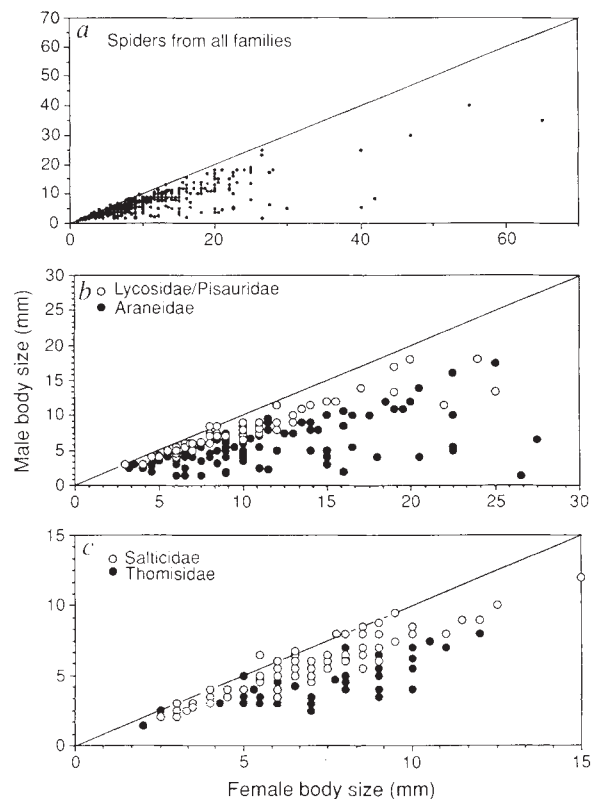


FIG. 3 *a*, A representative number of spiders from Western Europe^{30,31}, Japan³², Singapore³³ and Australia³⁴ (802 species in 31 families or sub-families). The line indicates equal body size, the regression was $y = 1.39 + 0.53x$, $r^2 = 0.75$. Further analysis regressing log (male body size) on to log (female body size) showed that overall the partial effect of hunting method (the effect once the effect of family size has been accounted for) was highly significant ($F_{1,779} = 45.29$): species that were sit-and-wait predators exhibited a higher degree of size dimorphism than species that were roving hunters. This is not affected by the average body size of species within the family; the interaction between body size and size dimorphism was not significant ($F_{1,767} = 1.87$). Nor was it much affected by phylogenetic relationships. Of the 17 major spider families, 7 were mainly hunters (Pisauridae, Lycosidae, Salticidae, Philodromidae, Sparassidae, Clubionidae and Gnaphosidae), 9 were web-builders (Uloboridae, Dictynidae, Linyphiidae, Araneidae, Metinae, Tetragnathinae, Theridiidae, Agelenidae and Atypidae) and in 1 (Thomisidae) most species are ambushers. The number of species that featured dwarfs (males of half the female length or smaller) as opposed to equally sized partners was 0:41 in the hunters, 93:66 in the web-builders and 17:3 in the ambushers. *b*, *c*, Scattergrams of male (ordinate) and female (abscissa) body size of active hunters (Lycosidae/Pisauridae, wolfspiders; Salticidae, jumping spiders) and sit-and-wait predators (Araneidae, orb spiders; Thomisidae crab spiders). The four families were chosen for their similarity in female body size and difference in hunting manner.

reproductive success per breeding) is maximized²³. Female reproductive success, R_f , increases linearly with body size (Fig. 4b), as is typical in invertebrates including spiders¹⁵. Male reproductive success R_m , is also assumed to increase linearly with size, but more steeply as sexual competition increases (Fig. 4b). For each sex, the solutions are equivalent²³

$$\begin{aligned} \partial R_f^* / \partial t_f &= d_{jf} R_f^* \text{ for females;} \\ \partial R_m^* / \partial t_m &= d_{jm} R_m^* \text{ for males} \end{aligned} \quad (1)$$

that is, at t_m^* , t_f^* , marginal gains in reproductive success through growth must be balanced by marginal losses to reproduction

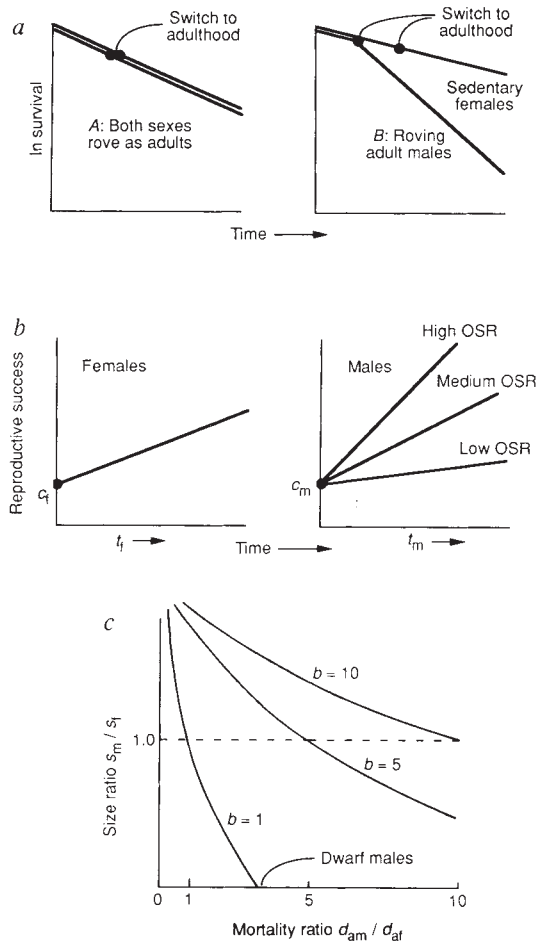


FIG. 4 Evolution of sexual dimorphism with differential adult mortality of the sexes. Mortality depends on mobility. *a*, The two spider life history patterns. In type A, both sexes rove throughout life, with similar (high) mortality in juvenile and adult stages. In type B, juveniles of both sexes are sedentary and thus have similar (low) mortality, but as adults, males rove and females remain sedentary and cryptic so that mortality is higher in males. *b*, the relationship between adult body size (s_m , s_f for males and females) and reproductive success. Adult body size is a function of age at maturity (t_m , t_f). Female reproductive success is $R_f = c_f + v_f s_f(t_f)$; and c_f , v_f are positive constants. Male reproductive success is assumed to increase more with size as operational sex ratio (OSR; β) increases: we assume that $R_m = c_m + v_m \beta s_m(t_m)$. Growth of sex i is assumed to be linear so that $s_i(t_i) = K_i t_i$, where K_i is the growth rate. *c*, Sexual size dimorphism. This is calculated from equations (1), using the above assumptions about mortality, growth and reproductive success. The optimal time of maturity in females is $t_f^* = (1/d_f) - c_f/(v_f K_f)$. Optimal time of male maturity is solved by iteration of $t_m^* = (1/d_m) - c_m/(\beta(t_m^*) v_m K_m)$, where $\beta(t_m^*)$ follows equation (2). If juvenile growth and mortality is similar in each sex ($d_{jm} = d_{jf} = d_j$; $K_m = K_f = K$), the ratio of male size to female size is then

$$s_m^*/s_f^* = [1 - (d_j c_m)/(K v_m) \beta(t_m^*)^{-1}] / [1 - (d_j c_f)/(K v_f)]$$

Numerical results of this solution are shown with $d_j = 0.1$; $K = 0.2$; $v_m = v_f = c_m = c_f = 1$. Size ratio is plotted against the adult mortality ratio at different values of b in equation (2).

through mortality. Thus juvenile mortality readily influences size by tuning the timing of maturity. Adult mortality exerts no influence, except (in the present model) on males, where R_m is a function of operational sex ratio, which depends on differential adult mortality. Assuming a primary sex ratio (males/females) of unity^{24,25} and that mortality is random and dependent only on mobility pattern, operational sex ratio is

$$\beta = b \left[\frac{d_{af}}{d_{am}} \exp(d_{jf} t_f^* - d_{jm} t_m^*) \right] \quad (2)$$

where b is the ratio: time per adult male seeking sexual encounter/time per female.

Assumptions (relevant to spiders) about growth and reproduction are given in Fig. 4b. Predictions of sexual dimorphism under differential adult mortality are shown in Fig. 4c, for different values of the constant b . If $b = 1$, both sexes spend equal proportions of adult life waiting for a mate; if $b > 1$ males spend a greater proportion of adult life seeking females than females spend seeking males, and *vice versa* for $b < 1$. Typically, $b \gg 1$, as each male spends most of its time seeking females, but females have a high 'time out' from mate seeking because of their relatively high parental investment²⁵. But in *Nephila* and many other spiders, the value of b will be low as once a male has located a female's web, he stays there. Figure 4c shows that relatively high mortality rates of adult males, coupled with relatively low values of b , favour production of small or dwarf males.

Our hypothesis predicts that, given a primary sex ratio of unity, the smaller sex will more frequently reach maturation but, because of higher adult mortality, may be less numerous in the adult population. The predictions are supported by evidence⁹ that male dwarfism is an evolutionary strategy in solitary, sedentary animals, that it evolves in low population densities (high search costs), and that it correlates with a female-biased adult sex ratio (reduced male competition). Our theory states that sex differences in life history trade-offs produce, and are produced by, biased adult and operational sex ratios. Higher male mortality reduces male competition, which favours earlier male maturation, and this in turn increases the proportion of adult males:females, which increases competition until the effects balance out: a form of selective feedback which tends to buffer skews in adult sex ratio at the expense of male body size. In *Nephila*, increased male mortality during the adult search phase is almost exactly counteracted by reduction in juvenile male growth stages.

Other examples of sedentary females with dwarf males may be found in barnacles (Crustacea, Cirripedia)^{5,6}, the hyperparasitic crab *Danalia curvata* (Crustacea, Peracaridae)^{26,27} and the angler fish *Ceratias* (Teleostei, Ceratiidae)^{7,28,29}. Indeed, it seems usual for bisexual groups with sedentary females to evolve dwarf males. Such an adult sex-ratio buffer as proposed here is quite distinct from Fisher's principle¹⁰ for sex ratio, which would operate only up to the end of parental care²¹. Provided that the advantage to males of large size is sufficiently low, there is nothing to prevent 'over-production' of adult males whenever early male maturity is not balanced by high adult male mortality. The eventual sex ratio can vary widely and our mechanism merely buffers its range. □

Received 15 May; accepted 24 September 1992.

1. Cunningham, J. T. *Sexual Dimorphism in the Animal Kingdom* (Black, London, 1900).
2. Hendrick, A. V. & Temeles, E. J. *Trends Ecol. Evol.* **4**, 136-139 (1989).
3. Gerhardt, U. *Zool. Anz.* **86**, 80-82 (1930).
4. Elgar, M. A., Ghaffar, N. & Read, A. J. *Zool.* **222**, 455-470 (1990).
5. Darwin, C. *A Monograph on the Cirripedia* (Royal Society, London, 1851).
6. Kaestner, A. *Invertebrate Zoology Vol. 2* (Wiley, New York, 1967).
7. Regan, C. T. *Proc. R. Soc.* **B97**, 386-399 (1925).
8. Darwin, C. *The Descent of Man and Selection in Relation to Sex* 2nd edn (Murray, London, 1894).
9. Ghiselin, M. T. *The Economy of Nature and the Evolution of Sex* (University of California Press, Berkeley, 1974).
10. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).

11. Hamilton, W. D. *Science* **156**, 477–488 (1967).
12. Vollrath, F. Z. *Tierpsychol.* **53**, 61–78 (1980).
13. Pagel, M. D. & Harvey, P. H. *Q. Rev. Biol.* **63**, 413–440 (1988).
14. Christenson, T. E. in *Contemporary Issues in Comparative Psychology* (ed. Dewsbury, D. A.) 149–174 (Sinauer, Sunderland, Massachusetts, 1990).
15. Vollrath, F. in *Ecophysiology of Spiders* (ed. Nentwig, W.) 357–370 (Springer, Berlin, 1987).
16. Newman, J. A. & Elgar, M. A. *Am. Nat.* **138**, 1372–1395 (1991).
17. Gertsch, W. J. *American Spiders* (Van Nostrand, New York, 1949).
18. Levi, H. W. J. *New York ent. Soc.* **58**, 59–91 (1955).
19. Watson, P. J. *Behav. Ecol. Sociobiol.* **26**, 77–90 (1990).
20. Gunnarson, B. J. *Zool.* **217**, 1–7 (1989).
21. Clutton-Brock, T. *The Evolution of Parental Care* (Princeton University Press, 1991).
22. Emlen, S. T. & Oring, L. W. *Science* **197**, 215–223 (1977).
23. Parker, G. A. J. *Fish. Biol.* (in the press).
24. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, 1982).
25. Clutton-Brock, T. & Parker, G. A. *Q. Rev. Biol.* (in the press).
26. Smith, G. *Fauna Flora Golf. Neapel* **29**, 1–123 (1906).
27. Caullery, M. *Mitt. zool. Stn Neapel* **18**, 583–643 (1908).
28. Bertelsen, E. *Dana Rep.* **39**, 1–176 (1951).
29. Beebe, W. *Zoologica* **6**, 149–241 (1934).
30. Lockett, G. H. & Millidge, A. F. *British Spiders* Vol. 1 (Royal Society, London, 1951).
31. Lockett, G. H. & Millidge, A. F. *British Spiders* (Royal Society, London, 1953).
32. Shinkai, E. & Takano, S. *Spiders of Japan* (Tokai University Press, 1984).
33. Koh, J. K. H. *A Guide to Common Singapore Spiders* (Singapore Science Centre, 1989).
34. Mascord, R. *Australian Spiders* (Reed, Wellington, New Zealand, 1970).

ACKNOWLEDGEMENTS. We thank S. Stearns for assistance.

Motor learning in a recurrent network model based on the vestibulo-ocular reflex

S. G. Lisberger* & T. J. Sejnowski†

* Department of Physiology, W. M. Keck Foundation Center for Integrative Neuroscience, and Neuroscience Graduate Program, University of California, San Francisco, California 94143, USA

† Salk Institute for Biological Science, Howard Hughes Medical Institute, 10010 N. Torrey Pines Road, La Jolla, California 92093, and Department of Biology, University of California, San Diego, California 92093, USA

MOST models of neural networks have assumed that neurons process information on a timescale of milliseconds and that the long-term modification of synaptic strengths underlies learning and memory¹. But neurons also have cellular mechanisms that operate on a timescale of tens or hundreds of milliseconds, such as a gradual rise in firing rate in response to injection of constant current² or a rapid rise followed by a slower adaptation³. These dynamic properties of neuronal responses are mediated by ion channels that are subject to modulation⁴. We demonstrate here how a neural network with recurrent feedback connections can convert long-term modulation of neural responses that occur over these intermediate timescales into changes in the amplitude of the steady output from the system. This general principle may be relevant to many feedback systems in the brain. Here it is applied to the vestibulo-ocular reflex, whose amplitude is subject to long-term adaptive modification by visual inputs⁵. The model reconciles apparently contradictory data on the neural locus of the cellular mechanisms that mediate this simple form of learning and memory.

We used model neurons in which the relationship between the summed inputs (V_{in}) and the firing rate output (V_{out}) was described by a single time constant (τ) which determined the rate at which V_{out} would rise in response to a stimulus increasing from zero to V_{in} . The input (V_{in}) to the model neuron was in the form of a brief ramp from zero to one unit and caused an output that rose more gradually from zero to one. The rising phase of the model neuron's output was a smoothed version of the input waveform when the time constant was 20 ms and a slower and smoothed version of the waveform when the time constant was 70 ms (Fig. 1). It is important to emphasize that the time constant in our model neurons is not equivalent to the membrane time constant; rather it describes the time course of changes in firing rate for a step change in input current and, in real neurons, it would be determined by the combined properties of a number of cellular mechanisms.

Figure 2a shows an example of a model network of neurons that can convert the small and transient difference between the two outputs of the model neuron in Fig. 1 into a large and sustained change in the steady-state output from the system. The input to the network is V and the output is E . The time courses of the responses of units T and F were determined by single time constants, like the model neuron in Fig. 1. Units P and B did not have time constants and their outputs at each time were equal to the sum of their inputs at the time. The strength of transmission of signals from one unit to the next was regulated by multiplication factors that scaled each input. For example, W_P scales the inputs from unit T to unit P and W_B scales the input from V to unit B. Other multiplication factors had values of +1 for excitatory connections and -1 for inhibitory connections.

Figure 2b shows how the network in Fig. 2a responded before and after the time constant of unit T was changed. To obtain the traces labelled 'Before', the values of W_B and W_P were both 1 and the time constants were 70 ms in units T and F. The input applied (V) then caused an identical output (E): the gain of the system, defined as E/V , was 1. When the time constant of unit T was decreased from 70 ms (Before) to 20 ms (traces labelled 'After'), the output of the network (E) was smaller and unit P showed a large response to the input waveform. Thus, changes in the time course of the response of unit T cause the steady-state gain of the system to be reduced, even though no changes were made to the scaling factors in the model.

Recurrent connections through units P, B, and F are the critical design feature that determines how the model works. These connections form a positive-feedback loop that allows a change in the time course of the output from one unit to be converted to a change in steady-state gain of the system. This can be understood intuitively by realizing that positive feedback will perform mathematical integration in a model that processes time-varying inputs. When the time constants in units T and F are the same and the values of W_P and W_B are 1, the two inputs to unit P cancel each other at all times and there is no input to be integrated. When the time constants in units T and F are

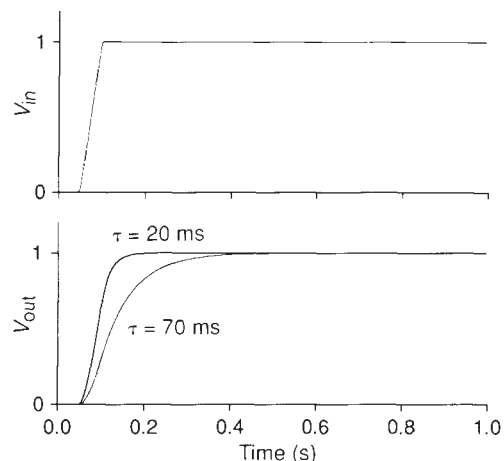


FIG. 1 Time-course of response of a simple model neuron in which a step input is transformed into an exponential rise in firing rate. The input (V_{in}) and output (V_{out}) of the unit are plotted as a function of time, showing the output when the time constant of the unit is 20 and 70 ms.

METHODS. Simulations were conducted with 'A Simulation Package', written by L. Optican and H. Goldstein. The models were described as block diagrams and the dynamic units were represented in Laplace notation as $1/(s\tau + 1)$, where τ is the time constant of filtering and s is the Laplace operator. The block diagrams were described in a Block Oriented Modeling Language (BOMOL) and a lexical analyser was used to convert the BOMOL code into the 'C' programming language. The program was then compiled and linked into a simulation shell that allowed control over the values of the parameters of the model and provided stimulus generation and graphical display of outputs. The time-step of simulation was 1 ms, much shorter than any of the time constants used here.

different, the two inputs to unit P differ transiently. This injects a transient into the positive-feedback loop. The transient is integrated and remembered and is expressed as a steady-state decrease in the output from the model.

The model network in Fig. 2a suggests a general algorithm for learning and may be applicable to the operation of many brain pathways that include recurrent connections. The network was selected for our studies specifically because of its relevance to the neural pathways that generate the vestibulo-ocular reflex (VOR)⁶. The VOR uses vestibular inputs to the brainstem and cerebellum to generate smooth eye movements that help stabilize images on the retina during head turns⁵. In Fig. 2, unit P represents a specific group of Purkinje cells in the cerebellum⁶ and unit B represents neurons in the brainstem⁷. The model must contain a positive-feedback pathway through units B, F and P because of evidence that such a recurrent pathway exists in the brain⁸. In the brain, the positive-feedback pathway carries an efference copy of eye velocity commands back to the cerebellum and is important in the generation of visually guided smooth

pursuit eye movements. During pursuit, the positive-feedback pathway is used to integrate transient visual motion inputs into a sustained eye velocity output⁹.

The gain of the VOR, defined as eye speed divided by head speed, is normally near 1, even in darkness when there is no possibility of visual correction¹⁰. The excellent performance of the VOR is established and maintained by a learning mechanism that uses the association of visual and vestibular inputs to guide adaptive changes in the gain of the VOR^{11,12}. Previous hypotheses have suggested that learning is mediated by changes in the gain of steady-state transmission at different synaptic relays but have disagreed about the loci of learning. Recordings under some conditions^{6,14} have been consistent with the suggestion¹³ that decreases in the gain of the VOR are mediated by increases in the value of W_P . Decreases in the gain of the VOR caused the output of the Purkinje cells, represented by unit P, to become modulated in phase with their vestibular inputs during the VOR. In contrast, recordings under other conditions^{6,7} appeared to contradict the hypothesis¹³: decreases in the gain of the VOR were associated with decreases in the value of W_P . It was postulated that decreases in the gain of the VOR would be mediated by parallel decreases in the value of W_B and W_P ⁶.

When applied to the VOR, the model in Fig. 2a offers a way to resolve the controversy between these two hypotheses as well as the apparently contradictory data^{6,14}. Analytical solution of the model in Fig. 2a shows that the model is stable only if W_B and W_P have the same value. Under these conditions, the gain of the VOR will be $W_P(\tau_T/\tau_F)$, where τ_T and τ_F are the time constants of units T and F, respectively. It was found that the neural equivalent of W_P was reduced to 81% of normal when the gain of the VOR was 0.18. Therefore, the solution of our model predicts that τ_T/τ_F should be 0.22. If we assume further that τ_F is fixed at 70 ms, then τ_T should be reduced to 18 ms. The analytical solution also reveals that the steady-state output of unit P during the VOR will be equal to $VW_P(\tau_F - \tau_T)/\tau_F$. Thus, the model predicts that the output of unit P will be in phase with vestibular input V when the gain of the VOR is low, as found by experiment^{6,7,14}.

Application of the model in Fig. 2a to the VOR therefore simulates motor learning in the VOR without contradicting any available data¹⁵. In addition, the output from the model (E in Fig. 2b) reproduces the observation that reductions in the gain of the VOR cause the eye velocity evoked by rapid head turns to show a transient overshoot before settling to a steady, sustained level¹⁶. In the model, motor learning is accomplished by several mechanisms at two or more loci. This suggests that learning in monkeys may be effected by a combination of changes in steady-state synaptic transmission in the brain stem and cerebellum and changes in time-course of the vestibular inputs to the cerebellum.

We described the time course of the response of model neurons to a step input with a single time constant so that we could both simulate (Fig. 2b) and calculate the performance of the network in Fig. 2a. Our simulations showed that changes in the transient component of a neuron's responses can be transformed into changes in the steady-state output from a network. Although a more complicated model would be required to describe the time-course of the responses of real neurons to a step increase in input current, the basic effects of changes in the transient response would be the same as those revealed here. In the brain, changes in the dynamics of neuronal responses could result from modulation of intrinsic cellular mechanisms, modulation of the strengths of local feedback connections, or changes in the selection of inputs with different dynamics. This third possibility is especially intriguing in a sensory-motor system like the VOR. In the vestibular system, some afferents respond to a head velocity stimulus with a tonic change in firing rate while others have a phasic/tonic response¹⁶. The mechanism demonstrated here could reduce the steady-state gain of the VOR by using local learning mechanisms to increase the

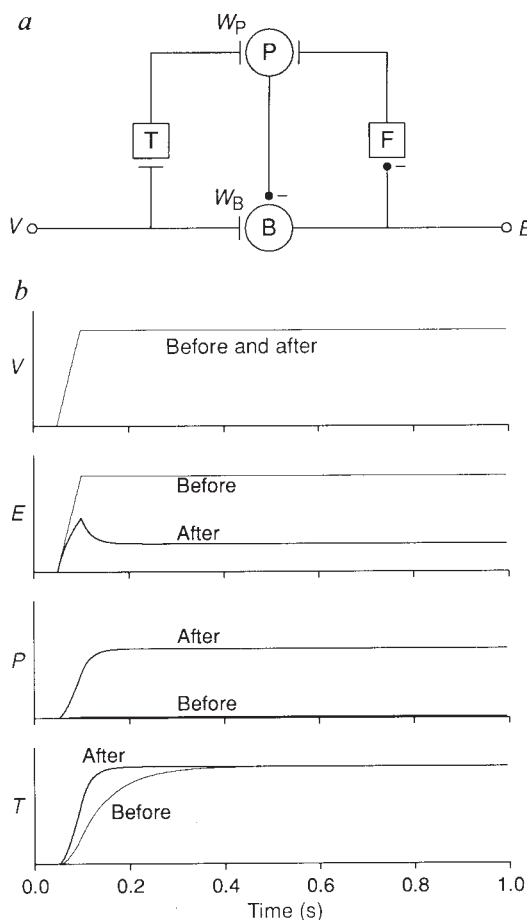


FIG. 2 Performance of a recurrent network that is based on the flow of signals in the primate vestibulo-ocular reflex. a, Diagram of the network, which includes a feedback loop with positive gain. Units T and F have the properties described in Fig. 1 and P and B operate as simple addition units without dynamical properties. Initially, units T and F both had time constants of 70 ms. W_P and W_B indicate the weights of the connections from unit T to unit P and from V to unit B, respectively. The connections indicated by minus signs have weights of -1 and the other connections have weights of $+1$. b, Effect of changing the time constant of unit T from 70 ms (Before) to 20 ms (After). Even though this did not change the steady-state response of unit T, it caused a profound reduction in the steady-state output from the network and a pronounced increase in the response of unit P. The steady-state gain of the system was 1.0 when τ_T was 70 ms and was 0.28 when τ_T was 20 ms. In applying the model to the VOR, unit P represented the 'horizontal gaze-velocity Purkinje cells' in the flocculus and ventral paraflocculus of the cerebellum^{5,7,8,14} and unit B represented the 'flocculus target neurons' in the vestibular nuclei⁷.

influence of the vestibular inputs that transmit phasic/tonic signals to the cerebellum while decreasing the influence of the inputs that transmit purely tonic signals¹⁵.

The recurrent positive feedback pathway in our neural network was critical for converting a subtle change in the time-course of neuronal responses into a change in the steady-state output of the system. Feedback connections are a general architectural feature of the brain and are found at many different levels, including the inputs and outputs of the cerebellum and the connections between areas of the cerebral cortex. The importance of recurrent connections in the pathways that mediate the VOR⁹ allows us to model how recurrent connections could contribute to learning and memory. The model raises the possibility that subtle changes in the function of individual cellular mechanisms may have profound effects on the output from specific behavioural systems and emphasizes the importance of understanding the architecture of the neural networks that convert cellular changes into changes in behavioural output. □

Received 5 May; accepted 18 September 1992.

- Churchland, P. S. & Sejnowski, T. J. *The Computational Brain* (MIT Press, Cambridge, MA, 1992).
- Storm, J. F. *Nature* **336**, 379–381 (1988).
- Bradley, K. & Somjen, G. G. *J. Physiol., Lond.* **156**, 75–92 (1961).
- Strong, J. A. & Kaczmarek, L. K. in *Neuromodulation: The Biochemical Control of Neuronal Excitability* (eds Kaczmarek, L. K. & Levitan, I. B.) (Oxford University Press, 1987).
- Lisberger, S. G. *Science* **242**, 728–735 (1988).
- Miles, F. A., Braitman, D. J. & Dow, B. M. *J. Neurophysiol.* **43**, 1477–1493 (1980).
- Lisberger, S. G. & Pavelko, T. A. *Science* **242**, 771–773 (1988).
- Miles, F. A., Fuller, J. H., Braitman, D. J. & Dow, B. M. *J. Neurophysiol.* **43**, 1437–1476 (1980).
- Lisberger, S. G., Morris, E. J. & Tyghsen, L. A. *Rev. Neurosci.* **10**, 97–129 (1987).
- Fuchs, A. F. & Kimm, J. J. *J. Neurophysiol.* **38**, 1140–1161 (1975).
- Gonshor, A. & Melvill Jones, G. *J. Physiol., Lond.* **256**, 381–414 (1976).
- Miles, F. A. & Lisberger, S. G. *Rev. Neurosci.* **4**, 273–299 (1981).
- Ito, M. *Brain Res.* **40**, 81–84 (1972).
- Watanabe, E. *Brain Res.* **297**, 169–174 (1984).
- Lisberger, S. G. & Sejnowski, T. J. *Univ. California San Diego Inst. Neural Comput. Tech. Rep.* 9201 (1992).
- Lisberger, S. G. & Pavelko, T. A. *J. Neurosci.* **6**, 346–354 (1986).

ACKNOWLEDGEMENTS. We thank numerous colleagues for their helpful comments. Research was supported by a grant from the Defense Advanced Research Project Agency, awarded through the Office of Naval Research.

A single amino-acid difference confers major pharmacological variation between human and rodent 5-HT_{1B} receptors

Donna Oksenberg^{*†}, Scot A. Marsters^{*}, Brian F. O'Dowd[‡], Hui Jin[‡], Sona Havlik[‡], Stephen J. Peroutka^{‡§} & Avi Ashkenazi^{*}

Departments of ^{*}Pulmonary Research and Gene Therapy, and [§]Neuroscience, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

[†]Department of Neurology, Stanford University School of Medicine, Stanford, California 94305, USA

[‡]Department of Pharmacology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

NEUROPSYCHIATRIC disorders such as anxiety, depression, migraine, vasospasm and epilepsy may involve different subtypes of the 5-hydroxytryptamine (5-HT) receptor^{1,2}. The 1B subtype, which has a unique pharmacology, was first identified in rodent brain^{3–7}. But a similar receptor could not be detected in human brain⁶, suggesting the absence in man of a receptor with equivalent function. Recently a human receptor gene was isolated (designated 5-HT_{1B} receptor^{8,9}, 5-HT_{1DB} receptor^{10,11}, or S12 receptor¹²) which shares 93% identity of the deduced protein sequence with rodent 5-HT_{1B} receptors^{13–15}. Although this receptor is identical

TABLE 1 Ligand-binding properties of the wild-type and T355N mutant human receptors compared with those of the rat and mouse 5-HT_{1B} receptors

Ligand	K _i (nM)			
	Human (wild type)	Human (T355N)	Rat ¹⁴ (5-HT _{1B})	Mouse ¹⁵ (5-HT _{1B})
Serotonergic				
5-HT	10 ± 1	8 ± 1	16 ± 1	39
5-CT	4 ± 1	3 ± 1	7 ± 1	10
DHE	6 ± 1	2 ± 1	4 ± 2	NA
RU24969	44 ± 4	2 ± 1	2 ± 1	10
Metergoline	25 ± 3	200 ± 40	129 ± 33	NA
Sumatriptan	38 ± 3	560 ± 100	465 ± 85	NA
Methysergide	130 ± 7	970 ± 130	1,823 ± 297	NA
8-OH-DPAT	1,600 ± 100	25,000 ± 1,000	>10,000	30,000
Methiothepin	12 ± 1	38 ± 8	13 ± 4	NA
β-adrenergic				
(–)Propranolol	8,100 ± 400	17 ± 1	57 ± 4	NA
(–)Pindolol	11,000 ± 1,000	20 ± 3	153 ± 62	69
(–)Alprenolol	11,000 ± 800	13 ± 1	NA	NA

The values are depicted as mean ± s.e.m. from 4 (wild type) and 3 (T355N) independent experiments done in triplicates. NA, data not available. Complementary DNAs encoding the wild-type and T355N mutant human receptors were inserted into the mammalian expression vector pRK5 and introduced by transient transfection into the human embryonic kidney 293 cell line by a modified calcium phosphate precipitation method²³. The cells were collected by centrifugation 48 h after transfection, lysed in ice-cold buffer (50 mM Tris-HCl, pH 7.4, containing 5 mM EDTA), homogenized, and sonicated for 10 s. Nuclei and intact cells were removed by centrifugation at 1,000g for 10 min. The supernatant was spun at 35,000g for 30 min and the resulting pellet, containing the microsomal membrane fraction, was resuspended in binding buffer containing 50 mM Tris-HCl (pH 7.4), 4 mM CaCl₂, 0.1% ascorbic acid, 10 mM pargyline and 1 μM leupeptine. Microsomal membranes (50 μg protein) were incubated with the ligands in binding buffer (30 min, 25 °C). Binding was terminated by the addition of 5 ml ice-cold 50 mM Tris-HCl (pH 7.8), rapid vacuum filtration through glass fibre filters, and two subsequent 5-ml washes. Specific binding was defined as the excess over blanks taken in the presence of 10^{–6} M cold 5-HT. Scatchard analyses of saturation binding of [³H]5-HT showed two populations of binding sites; equilibrium dissociation constants (K_D) for the wild-type and mutant receptor, respectively, were 4.6 ± 1.4 and 3.9 ± 1.2 nM (high-affinity sites); 72 ± 24 and 75 ± 14 nM (low-affinity sites). The respective receptor densities in pmol per gram of protein were 1,000 ± 320 and 1,440 ± 610 (high-affinity sites), 13,670 ± 1,860 and 25,330 ± 4,480 (low-affinity sites). Specific binding of [³H]5-HT was not detectable in untransfected cells (not shown), indicating that these cells do not express significant levels of endogenous 5-HT receptors. Equilibrium inhibition constants (K_i) were determined according to the following equation: K_i = IC₅₀/(1 + [T]/K_D), where IC₅₀ is the concentration of competing ligand required for 50% inhibition of [³H]5-HT binding, [T] is the concentration of the [³H]5-HT tracer (3 nM), and K_D is the high-affinity constant of [³H]5-HT, as determined by saturation binding. The data were analysed by nonlinear least-square fitting using the EBDA²⁴ and LIGAND²⁵ programs.

to rodent 5-HT_{1B} receptors in binding to 5-HT, it differs profoundly in binding to many drugs. Here we show that replacement of a single amino acid in the human receptor (threonine at residue 355) with a corresponding asparagine found in rodent 5-HT_{1B} receptors renders the pharmacology of the receptors essentially identical. This demonstrates that the human gene does indeed encode a 1B receptor, which is likely to have the same biological functions as the rodent 5-HT_{1B} receptor. In addition, these findings show that minute sequence differences between homologues of the same receptor from different species can cause large pharmacological variation. Thus, drug–receptor interactions should not be extrapolated from animal to human species without verification.

The human and rodent 5-HT_{1B} receptors bind to 5-HT with comparably high affinity. But the human receptor binds with much lower affinity to the serotonergic agonist RU 24969 and to the β-adrenergic receptor antagonists propranolol, pindolol and alprenolol, and with higher affinity to several serotonergic drugs, including sumatriptan and 8-hydroxy-2-(di-*n*-propyl-amino)tetralin (8-OH-DPAT)^{8–14} (Table 1).

The 32 amino-acid differences between the human and rat receptors are scattered throughout the molecule, but only eight are found in the transmembrane domains, which are thought to contain the ligand-binding pocket (Fig. 1). An asparagine residue in the seventh transmembrane segment has been implicated in β-antagonist binding to the β-adrenergic receptor¹⁶ and to human 5-HT_{1A} receptors¹⁷. Notably, an asparagine

FIG. 1 Seven-transmembrane-segment model of the human 5-HT_{1B} receptor. The amino acids found in the human receptor^{8,12} are circled; residues that are different or absent in the rat 5-HT_{1B} receptor^{13,14} are shown next to the corresponding positions. The rat sequence is 98% identical to the mouse¹⁵ 5-HT_{1B} receptor sequence. The dark box indicates position 355, in which the threonine of the human receptor was replaced with the corresponding asparagine found in the rat and the mouse 5-HT_{1B} receptors. METHODS. Threonine codon 355 of the human receptor cDNA⁸ was replaced with an asparagine codon by oligonucleotide-directed mutagenesis. A 39-base synthetic oligonucleotide carrying the asparagine codon (containing two mismatches with the threonine codon) (GTTGAGATAGCCAGCCAGTTGAA-GAAGTCAAAGATGGC) was used as the mutagenesis primer. This primer was hybridized to a single-stranded DNA containing the wild-type sequence and extended by T4 DNA polymerase, and the double-stranded DNA plasmid was introduced into *Escherichia coli*. Bacterial colonies containing the T355N mutation were identified by two consecutive hybridization analyses and a single clone carrying the mutation was selected and confirmed by DNA sequencing.

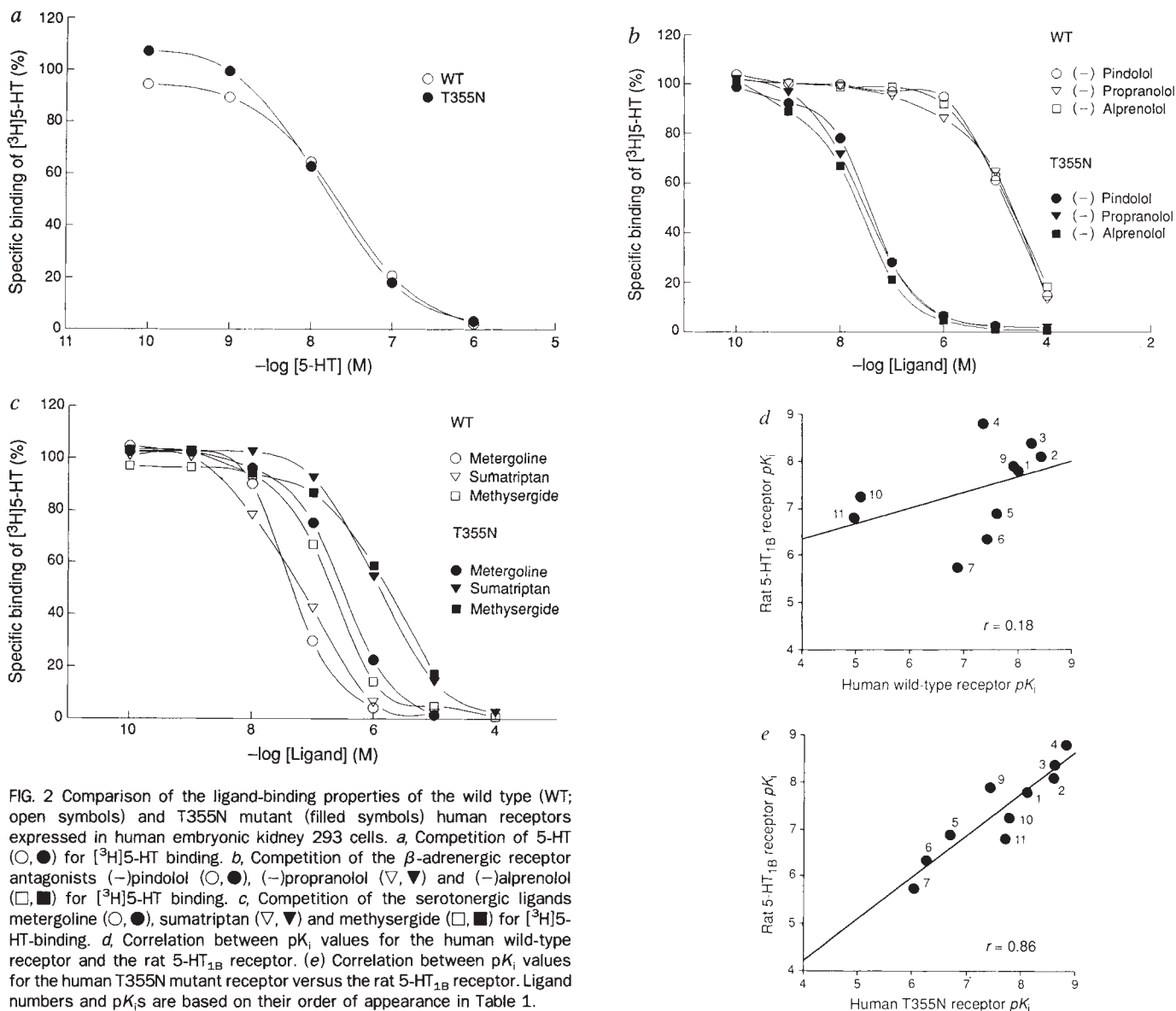
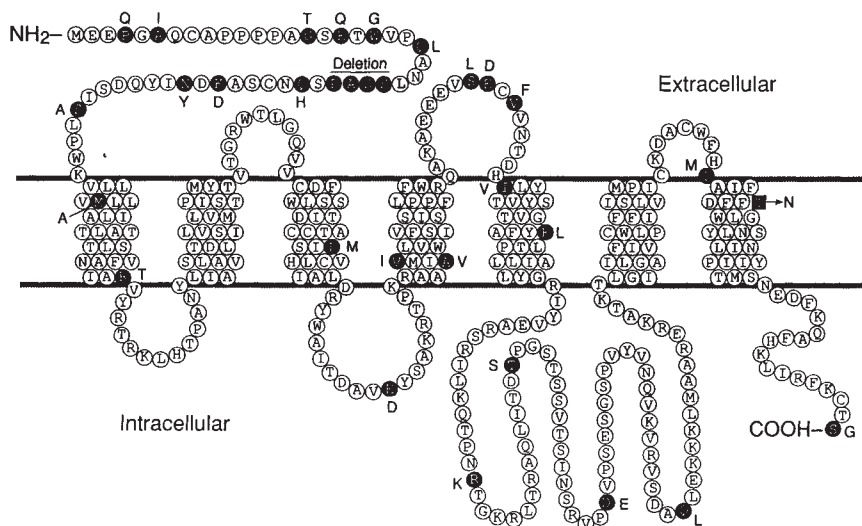


FIG. 2 Comparison of the ligand-binding properties of the wild type (WT; open symbols) and T355N mutant (filled symbols) human receptors expressed in human embryonic kidney 293 cells. *a*, Competition of 5-HT (○, ●) for [³H]5-HT binding. *b*, Competition of the β -adrenergic receptor antagonists (-)pindolol (○, ●), (-)propranolol (▽, ▼) and (-)alprenolol (□, ■) for [³H]5-HT binding. *c*, Competition of the serotonergic ligands metergoline (○, ●), sumatriptan (▽, ▼) and methysergide (□, ■) for [³H]5-HT-binding. *d*, Correlation between pK_i values for the human wild-type receptor and the rat 5-HT_{1B} receptor. *e*, Correlation between pK_i values for the human T355N mutant receptor versus the rat 5-HT_{1B} receptor. Ligand numbers and pK_is are based on their order of appearance in Table 1.

residue is found at this location in all cloned 5-HT receptors that bind β -antagonists; rat^{13,14} and mouse¹⁵ 5-HT_{1B} receptors, rat¹⁸ and human^{19,20} 5-HT_{1A} receptors, and in β -adrenergic receptors²¹. In contrast, a threonine residue is found at this position (355) in the human 5-HT_{1B} receptor⁸⁻¹². To assess the contribution of this threonine to the distinct pharmacology of the human receptor, we replaced it with the corresponding asparagine found in rodent 5-HT_{1B} receptors by site-directed mutagenesis (Fig. 1). Scatchard analysis showed that the wild-type and mutant (T355N) receptors bind [³H]5-HT with similar affinity, each with two binding-site populations (K_D = 4.6 and 72 nM, wild type; 3.9 and 75 nM, mutant; data not shown). Competition analysis also showed similar 5-HT affinities for the wild-type and mutant receptors (Fig. 2a; Table 1). Thus, the T355N substitution does not affect binding to the natural ligand 5-HT.

To compare the pharmacology of the wild-type and mutant human receptors with that of rodent 5-HT_{1B} receptors, we carried out competition assays with serotonergic and β -adrenergic drugs from various chemical classes. Like 5-HT, 5-carboxyamidotryptamine (5-CT) and dihydroergotamine (DHE), which have similar inhibition constants (K_i) with the wild-type human and the rodent receptors, showed comparable K_i s for T355N (Table 1). In contrast, the β -antagonists propranolol, pindolol and alprenolol had markedly reduced K_i s with T355N, which were comparable to the K_i s seen with the rodent receptors (Fig. 2b; Table 1). The K_i of the serotonergic ligand RU 24969 was ~20-fold lower with T355N than with the wild type, and similar to the K_i s for the rodent receptors (Table 1). Metergoline, sumatriptan, methysergide and 8-OH-DPAT bound to T355N with K_i s that were 7–15 times higher than those seen with the wild type, and more similar to K_i s observed for the rodent receptors (Fig. 2c; Table 1). Methiothepin, which has similar K_i s with the wild-type human and the rodent receptors, showed a fourfold higher K_i with T355N (Table 1). The K_i s of 11 ligands showed no correlation between the human wild type and T355N (r = 0.03), or the human wild-type and rat receptors (r = 0.18), but showed a significant correlation between T355N and the rat receptor (r = 0.86; P < 0.0025) (Fig. 2d, e).

Thus, substitution of threonine 355 of the human receptor with the corresponding asparagine of rodent 5-HT_{1B} receptors confers the pharmacology of a 1B subtype on the human receptor. Taken together with the high sequence homology between these receptors, this demonstrates that the 5-HT receptor gene investigated here does indeed encode the human species variant of the 1B 5-HT receptor subtype. Therefore, given that the human and rodent 5-HT_{1B} receptors bind the natural ligand, 5-HT, with the same affinity, it is likely that despite their dramatically different pharmacology, these receptors are equivalent in function.

The identification of a human 5-HT_{1B} receptor is of biological and clinical significance. Whereas this receptor and the human 5-HT_{1D} receptor²² share only 68% amino-acid sequence identity and are encoded on different chromosomes, their pharmacological properties are virtually indistinguishable⁸⁻¹². Therefore, biological functions such as appetite control, migraine reduction and regulation of 5-HT release, attributed previously to the 5-HT_{1D} receptor on the basis of drugs presumed to be selective for this subtype, indeed may be mediated by either the 5-HT_{1D} or the 5-HT_{1B} receptor, or by both. Understanding the individual functions of these receptors thus may help improve the clinical application of compounds such as the antimigraine drug sumatriptan, which binds to both the 5-HT_{1D} and 5-HT_{1B} receptors.

Finally, our results show that a single amino-acid difference can cause dramatic pharmacological variation between species homologues of the same receptor. Thus, ligand-binding properties of a given receptor cannot be extrapolated across species without independent validation, even in the context of very high interspecies sequence identity. This has important implications

for the design and therapeutic use of receptor-selective drugs when these are based on measurements of selectivity and efficacy in non-human species. □

Received 19 June; accepted 28 September 1992.

- Hartig, P. R. et al. *Neuropsychopharmacology* **3**, 335–347 (1990).
- Peroutka, S. J. *Neuropharmacology* **31**, 609–613 (1992).
- Pedigo, N. W., Yamamura, H. I. & Nelson, D. L. *J. Neurochem.* **36**, 220–226 (1981).
- Schnellman, R. G., Waters, S. J. & Nelson, D. L. *J. Neurochem.* **42**, 65–70 (1984).
- Hoyer, D., Engel, G. & Kalkman, H. O. *Eur. J. Pharmac.* **118**, 1–12 (1985).
- Hoyer, D., Engel, G. & Kalkman, H. O. *Eur. J. Pharmac.* **118**, 13–23 (1985).
- Hoyer, D. & Middlemiss, D. N. *Trends pharmac. Sci.* **10**, 130–132 (1989).
- Jin, H. et al. *J. Biol. Chem.* **267**, 5735–5738 (1992).
- Hamblin, M. W., Metcalf, M. A., McGuffin, R. W. & Karpells, S. *Biochem. biophys. Res. Commun.* **184**, 752–759 (1992).
- Weinshank, R. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3630–3634 (1992).
- Demchishyn, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5522–5526 (1992).
- Olav-Levy, F. et al. *J. Biol. Chem.* **267**, 7553–7562 (1992).
- Voigt, M. M., Laurie, D. J., Seeburg, P. H. & Bach, A. *EMBO J.* **10**, 4017–4023 (1991).
- Adham, N. et al. *Molec. Pharmac.* **41**, 1–7 (1992).
- Maroteaux, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3020–3024 (1992).
- Suryanarayana, A., Daunt, D. A., VonZastrow, M. & Kobilka, B. K. *J. Biol. Chem.* **266**, 15488–15492 (1991).
- Guan, X. G., Peroutka, S. J. & Kobilka, B. K. *Molec. Pharmac.* **41**, 695–698 (1992).
- Albert, P. R. et al. *J. Biol. Chem.* **265**, 5825–5832 (1990).
- Kobilka, B. K. et al. *Nature* **329**, 75–79 (1987).
- Fargin, A. et al. *J. Biol. Chem.* **264**, 4848–4852 (1989).
- Findlay, J. & Eliopoulos, E. *Trends pharmac. Sci.* **11**, 492–499 (1990).
- Hamblin, M. W. & Metcalf, M. A. *Molec. Pharmac.* **40**, 143–148 (1991).
- Ashkenazi, A., Peralta, E. G., Winslow, J. W., Ramchandran, J. & Capon, D. J. *Cell* **56**, 487–493.
- McPherson, G. A. *Comp. Prog. Biomed.* **17**, 107–114 (1983).
- Munson, P. J. & Rodbard, D. *Analyt. Biochem.* **107**, 220–239 (1989).

ACKNOWLEDGEMENTS. We thank C. Clark and R. Ward for comments on the manuscript. This work was supported by the NIH, the Kleiner Family Foundation and Genentech, Inc.

Positive feedback of glutamate exocytosis by metabotropic presynaptic receptor stimulation

Inmaculada Herrero, M. Teresa Miras-Portugal & José Sánchez-Prieto*

Departamento de Bioquímica, Facultad de Veterinaria, Universidad Complutense, 28040 Madrid, Spain

GLUTAMATE is important in several forms of synaptic plasticity such as long-term potentiation, and in neuronal cell degeneration^{1,2}. Glutamate activates several types of receptors, including a metabotropic receptor that is sensitive to *trans*-1-amino-cyclopentyl-1,3-dicarboxylate, coupled to G protein(s) and linked to inositol phospholipid metabolism^{3–6}. The activation of the metabotropic receptor in neurons generates inositol 1,4,5-trisphosphate, which causes the release of Ca²⁺ from intracellular stores and diacylglycerol, which activates protein kinase C^{7–9}. In nerve terminals, the activation of presynaptic protein kinase C with phorbol esters enhances glutamate release¹⁰. But the presynaptic receptor involved in this protein kinase C-mediated increase in the release of glutamate has not yet been identified. Here we demonstrate the presence of a presynaptic glutamate receptor of the metabotropic type that mediates an enhancement of glutamate exocytosis in cerebrocortical nerve terminals. Interestingly, this potentiation of glutamate release is observed only in the presence of arachidonic acid, which may reflect that this positive feedback control of glutamate exocytosis operates in concert with other pre- or post-synaptic events of the glutamatergic neurotransmission that generate arachidonic acid. This presynaptic glutamate receptor may have a physiological role in the maintenance of long-term potentiation where there is an increase in glutamate release mediated by postsynaptically generated arachidonic acid¹¹.

The addition of the selective agonist for the metabotropic receptor, (1S,3R)-1-aminocyclopentane-1,3-dicarboxylic acid

* To whom correspondence should be addressed.

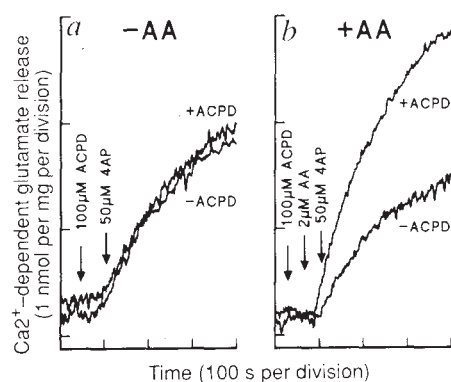


FIG. 1 ACPD enhances the 4-aminopyridine(4AP)-evoked release of glutamate in the presence of arachidonic acid (AA). *a*, Glutamate release evoked by 50 μ M 4-aminopyridine in the absence and *b*, in the presence of 2 μ M arachidonic acid. The Ca^{2+} -dependent glutamate release was calculated by subtracting release at a free Ca^{2+} concentration of 200 nM, from release at 1.33 mM CaCl_2 . Each trace is the mean of three independent experiments. METHODS. The P_2 fraction of synaptosomes was prepared from the cerebral cortices of male Wistar rats²⁸. Synaptosomes (1 mg pellets) were resuspended into 1.5 ml of incubation medium (122 mM NaCl, 3.1 mM KCl, 0.4 mM KH_2PO_4 , 5 mM NaHCO_3 , 1.2 mM MgSO_4 , 10 mM glucose and 20 mM TES buffer, pH 7.4) and incubated for 1 h at 37 $^\circ\text{C}$ in the presence of 16 μ M bovine serum albumin (BSA) that was essentially fatty-acid free (Sigma) to prevent the effects of fatty acids released from synaptosomes during the incubation²⁹. Glutamate release was determined as described^{30,31}. After preincubation for 1 h in the presence of BSA, the synaptosomes were pelleted and resuspended in fresh incubation medium without albumin. An aliquot (1 ml) was transferred to a stirred cuvette containing 1 mM nicotinamide adenine dinucleotide phosphate (NADP^+), 50 U glutamate dehydrogenase, and 1.33 mM CaCl_2 or 200 nM free Ca^{2+} (50 μ M EGTA and 38 μ M CaCl_2)³². Fluorescence was measured using a Perkin-Elmer model LS-50 luminescence spectrometer at 340 and 460 nm. Data were collected at 2-s intervals. Methyl arachidonate, 4 β -phorbol dibutyrate (4 β -PDBu) and 12-*O*-tetradecanoylphorbol-13-acetate were prepared in dimethylsulphoxide. Arachidonic acid was stored under nitrogen atmosphere at -70 $^\circ\text{C}$ as a 15 mM stock suspension in water or occasionally in dimethylsulphoxide. Before use, the stock suspension was thawed and sonicated for 1 min.

(1S,3R-ACPD), to cerebrocortical nerve terminals (synaptosomes) increased the glutamate release evoked by a submaximal concentration of 4-aminopyridine in the presence (Fig. 1*b*) but not in the absence of arachidonic acid (Fig. 1*a*). Control experiments showed that this increase in glutamate release caused by

FIG. 3 Pharmacological characterization of the ACPD-induced enhancement of glutamate release. In all experiments (including the control) arachidonic acid was added 30 s before depolarization of synaptosomes with 50 μ M 4-aminopyridine. Control glutamate release (100%) was 1.20 ± 0.11 nmol glutamate per mg of protein $\pm$ s.d. ($n=5$). *a*, Dose-response curve of ACPD on glutamate release. Different concentrations of ACPD were added 1 min before depolarization of synaptosomes with 50 μ M 4-aminopyridine. *b*, Effects of agonists and antagonists of excitatory amino-acid receptors and other treatments in the ACPD-induced potentiation of glutamate release. The glutamate receptor antagonists CNQX (5 μ M), L-AP3 (100 μ M) and AP5 (100 μ M) were added 2 min before depolarization of synaptosomes with 50 μ M 4-aminopyridine. The agonists ACPD (100 μ M), ibotenate (Ibo) (100 μ M) and quisqualate (Quis) (100 μ M) were added 1 min before depolarization. In experiments with pertussis toxin (PTX), synaptosomes were preincubated with the toxin (1.5 μ g per mg of protein) for 2 h. In experiments with staurosporine, the synaptosomes were incubated with protein kinase inhibitor at 100 nM for 30 min. A group of synaptosomes were preincubated with 1 μ M 12-*O*-tetradecanoylphorbol-13-acetate (TPA) for 30 min. After preincubation the

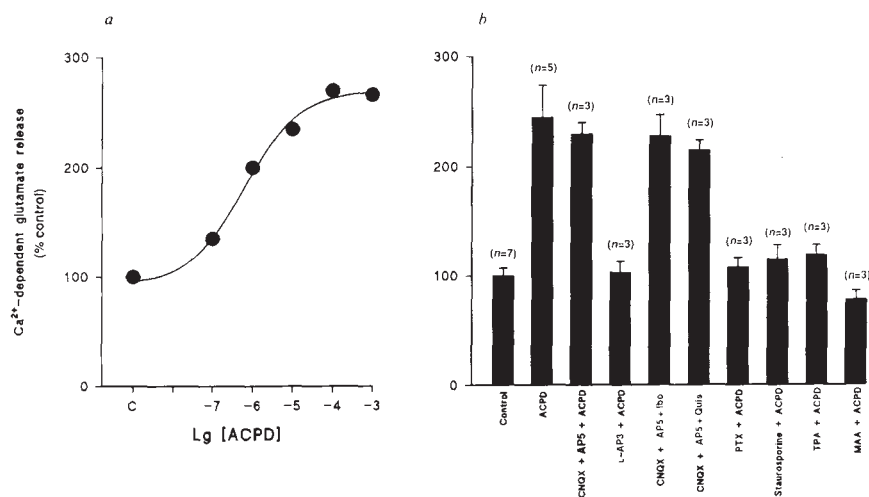


FIG. 2 Effects of arachidonic acid in the sensitivity of glutamate release to 4 β -PDBu (4 β -phorbol dibutyrate). Different concentrations of 4 β -PDBu and 2 μ M arachidonic acid (AA) or 2 μ M methyl arachidonate (MAA), were added 60 s and 30 s, respectively, before depolarization of synaptosomes with 1 mM 4-aminopyridine.

ACPD in the presence of arachidonic acid is not due to the inhibition by ACPD of the glutamate/aspartate carrier, as we did not detect any change in the external glutamate concentration of polarized synaptosomes even after a 30-min incubation in the presence of 100 μ M ACPD (data not shown).

The lack of effect of ACPD in the absence of arachidonic acid suggests a relatively low sensitivity of glutamate exocytosis to potentiation by phorbol esters (Fig. 2) which are known to mimic the physiological activator of protein kinase C (PKC), diacylglycerol¹³. In contrast, in the presence of arachidonic acid the sensitivity of glutamate exocytosis to phorbol esters is increased, probably through synergistic activation of PKC¹⁴ as this increase in sensitivity is not observed in the presence of methyl arachidonate, which does not activate PKC¹⁵.

In addition to the metabotropic receptor, glutamate activates at least three distinct ionotropic receptors: the *N*-methyl-D-aspartate (NMDA), the α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) and the kainate subtypes^{16,17}. It is unlikely that the potentiation of glutamate release observed in the presence of ACPD is mediated by the activation of ionotropic receptors. Firstly, ACPD was effective in potentiating glutamate exocytosis at low concentrations, with only 1×10^{-6} M

synaptosomes were pelleted and assayed for glutamate release. Experiments with methyl arachidonate (MAA) were similar to the control but used methyl arachidonate instead of arachidonic acid.

ACPD being required to produce a half-maximal response (EC_{50}) (Fig. 3a), a concentration far lower than those that affect ionotropic receptors^{4,18}. Secondly, the ACPD-mediated potentiation of glutamate release was still observed in the presence of the ionotropic receptor antagonists 6-cyano-7-nitroquinoxaline-2,3-dione, CNQX¹⁹ and 2-amino-5-phosphonovaleric acid, AP5²⁰ (Fig. 3b). Thirdly, ACPD-mediated potentiation of glutamate exocytosis was prevented by the relatively selective antagonist of the metabotropic receptor, L-2-amino-3-phosphonopropionate, L-AP3 (refs 4, 21) (Fig. 3b). Finally, other agonists of the metabotropic receptor such as quisqualate and ibotenate also potentiated glutamate release in the presence of the ionotropic receptor antagonists CNQX and AP5 (Fig. 3b).

Ionotropic glutamate receptors contain cation-specific ionic channels, whereas the metabotropic receptor is coupled to a G protein sensitive to pertussis toxin³. Pretreatment of synaptosomes with pertussis toxin had no effect in control release (data not shown) but blocked the stimulatory effect of ACPD on glutamate release (Fig. 3b).

The invasion of the nerve terminal by repetitive action potentials is mimicked by stimulation with 4-aminopyridine which evokes transmitter release in nerve terminals by inducing tetrodotoxin-sensitive transient depolarizations²². Activation of presynaptic PKC with phorbol esters enhances these depolarizations, causing a longer opening of the non-inactivating Ca^{2+} channels that are coupled to glutamate exocytosis²³ and thus an enhanced entry of Ca^{2+} and, consequently, an increased

release of glutamate¹⁰. We investigated the mechanism by which ACPD enhances glutamate exocytosis by monitoring the plasma membrane potential and the cytoplasmic free Ca^{2+} concentration in parallel. The results clearly show that the ACPD-induced enhancement in 4-aminopyridine-evoked glutamate release is consistent with an increased 4-aminopyridine-evoked depolarization and also an increased 4-aminopyridine-evoked elevation of cytoplasmic free Ca^{2+} (Fig. 4).

The fact that the metabotropic agonist ACPD mimics the effects of PKC activation of glutamate release, suggests that the actions of the metabotropic receptor activation on glutamate exocytosis are probably mediated by the diacylglycerol-PKC branch of the phosphoinositide cascade rather than by the generation of $InsP_3$. Consistent with this, we observed that the potentiating effect by ACPD and arachidonic acid on glutamate release was prevented both in synaptosomes preincubated with the protein kinase inhibitor staurosporine and in synaptosomes in which the PKC activity was 'downregulated' by pretreatment with a relatively high concentration of phorbol esters (Fig. 3b). Moreover, when methyl arachidonate, which does not activate PKC¹⁵, was used instead of arachidonic acid, the potentiating effect by the metabotropic agonist was also abolished (Fig. 3b). All of these results indicate that the effect of ACPD and arachidonic acid on glutamate release is mediated by the synergistic activation of PKC by arachidonic acid and by the diacylglycerol generated after metabotropic receptor activation.

We provide strong evidence for the existence of a metabotropic glutamate receptor located at the presynaptic site of the glutamatergic synapses. An interesting property of this presynaptic receptor is its involvement in the potentiation of glutamate exocytosis. Such a presynaptic control mechanism, which would be expected to result in neurotoxicity², may in fact underlie a physiologically relevant modulatory pathway, as it operates only in the presence of arachidonic acid. Long-term potentiation in the hippocampus is thought to be initiated by the activation of a postsynaptic NMDA receptor and the entry of calcium through its associated channel¹, whereas the maintenance of long-term potentiation is, at least in part, mediated by a sustained increase in glutamate release²⁴. The presynaptic mechanism for potentiation of glutamate release described here requires ACPD or glutamate and arachidonic acid. The necessity of glutamate suggests that this presynaptic control is restricted to recently active presynaptic nerve terminals. Moreover, the need for arachidonic acid may indicate that this presynaptic mechanism only operates in concert with postsynaptic events, such as the generation of arachidonic acid as a retrograde messenger²⁵. On the other hand, when only glutamate or arachidonic acid are present, the potentiation mechanism for glutamate release is inactive. In fact, we have reported that arachidonic acid alone depresses glutamate exocytosis by a mechanism that causes a reduction in both the depolarization and in the Ca^{2+} entry evoked by 4-aminopyridine²⁶; this mechanism is independent of the activation of PKC because this inhibitory action is also observed in the presence of methyl arachidonate²⁷, which does not activate PKC. □

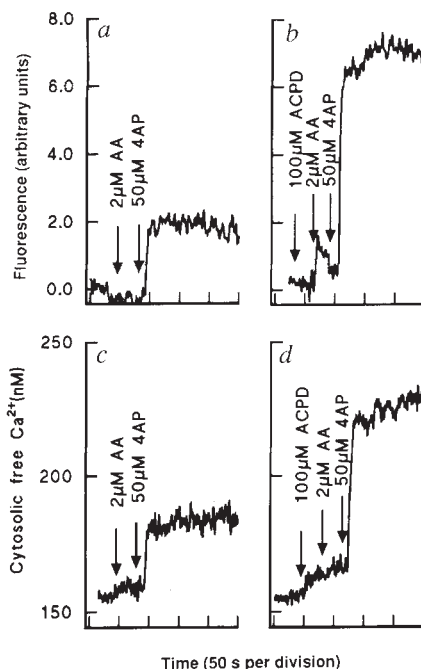


FIG. 4 Enhancement by ACPD of the 4-aminopyridine (4AP)-evoked depolarization and of the 4AP-evoked increase in cytoplasmic free Ca^{2+} concentration. Arachidonic acid (AA) at 2 μ M was added 30 s before depolarization of synaptosomes with 50 μ M 4AP. Results are the means of at least three independent experiments.

METHODS. The cytosolic free Ca^{2+} concentration was determined with Fura-2 (ref. 33). Synaptosomes (2 mg ml⁻¹) were resuspended in incubation medium with BSA. After 5 min, 1.33 mM $CaCl_2$ and 5 μ M Fura-2-acetoxymethyl ester were added and after 35 min the synaptosomes were pelleted and resuspended into fresh medium containing 1.33 mM $CaCl_2$ but no albumin. Fluorescence was monitored at 340 and 510 nm. Data were collected at 2-s intervals. Synaptosomal plasma membrane potential was monitored using a potential-sensitive cationic cyanine dye, 3,3'-diethyl-thiadicarbocyanine iodide (DiSC₂(5))³⁴. Synaptosomes in incubation medium with BSA were preincubated at 37 °C for 1 h and then pelleted and resuspended (1 mg ml⁻¹) in incubation medium without BSA but with 5 μ M DiSC₂(5) and 1.33 mM $CaCl_2$. After 5 min for equilibration, fluorescence was determined at 643 and 680 nm. Data were collected at 2-s intervals.

Received 9 June; accepted 17 September 1992.

- Collingridge, G. L. & Bliss, R. V. P. *Trends Neurosci.* **10**, 288–293 (1987).
- Choi, D. W. & Rothman, S. M. *A. Rev. Neurosci.* **13**, 171–182 (1990).
- Masu, M., Tanabe, Y., Tsushima, K., Shigemoto, R. & Nakanishi, S. *Nature* **349**, 760–765 (1991).
- Schoepp, D., Bockaert, J. & Sladeczek, F. *Trends Pharmacol. Sci.* **11**, 508–515 (1990).
- Sugiyama, H., Ito, I. & Hirono, C. *Nature* **325**, 531–533 (1987).
- Nicoletti, F., Wroblewski, J. T., Fadda, E. & Costa, E. *Neuropharmacology* **27**, 551–556 (1988).
- Sladeczek, F., Pin, J.-P., Recasens, M., Bockaert, J. & Weiss, S. *Nature* **317**, 717–719 (1985).
- Murphy, S. N. & Miller, J. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8737–8741 (1988).
- Manzoni, O. J. J. et al. *Neurosci. Lett.* **109**, 146–151 (1990).
- Barrie, A. P., Nicholls, D. G., Sánchez-Prieto, J. & Sihra, T. S. *J. Neurochem.* **57**, 1398–1404 (1991).
- Lynch, M. A., Errington, M. L., Bliss, T. V. P. *Neuroscience* **30**, 693–701 (1989).
- Palmer, E., Monaghan, D. T. & Cotman, C. W. *Eur. J. Pharmacol.* **166**, 585–587 (1989).
- Castagna, M. et al. *J. Biol. Chem.* **257**, 7847–7851 (1982).
- Shinomura, T., Asakura, Y., Yoshida, K. & Nishizuka, Y. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5149–5153 (1991).
- Murakami, K. & Routenberg, A. *FEBS Lett.* **192**, 189–193 (1985).
- Mayer, M. L. & Westbrook, G. L. *Prog. Neurobiol.* **28**, 197–276 (1990).
- Monaghan, D. T., Bridges, R. J. & Cotman, C. W. *A. Rev. Pharmacol. Toxicol.* **29**, 365–402 (1989).

18. Schoepp, D. D., Jonshon, B. G., Salhoff, C. R. & Hillman, C. C. *Soc. Neurosci. Abstr.* **16**, 618 (1990).
19. Honoré, T. *et al. Science* **241**, 701–703 (1988).
20. Davies, J. A., Francis, A., Jones, A. W. & Watkins, J. C. *Neurosci. Lett.* **21**, 77–82 (1981).
21. Schoepp, D. D. & Johnson, B. G. *J. Neurochem.* **53**, 273–278 (1989).
22. Tibbs, G. R., Barrie, A. P., Van-Miegheem, F., McMahon, H. T. & Nicholls, D. G. *J. Neurochem.* **53**, 1693–1699 (1989).
23. McMahon, H. T. & Nicholls, D. G. *J. Neurochem.* **56**, 86–94 (1991).
24. Bliss, T. V. P., Douglas, R. M., Errington, M. L. & Lynch, M. A. *J. Physiol., Lond.* **377**, 391–408 (1986).
25. Bliss, T. V. P. & Lynch, M. A. in *Neurology and Neurobiology* Vol. 35 (eds Landfield, P. W. & Deadwyler, S. W.) 3–72 (Liss, New York, 1988).
26. Herrero, I., Miras-Portugal, M. T. & Sánchez-Prieto, J. *J. Neurochem.* **57**, 718–721 (1991).
27. Herrero, I., Miras-Portugal, M. T. & Sánchez-Prieto, J. *FEBS Lett.* **296**, 317–319 (1992).
28. Nicholls, D. G. *Biochem. J.* **170**, 511–522 (1978).
29. Herrero, I., Castro, E., Miras-Portugal, M. T. & Sánchez-Prieto, J. *Neurosci. Lett.* **126**, 41–44 (1991).
30. Rubio, I., Torres, M., Miras-Portugal, M. T. & Sánchez-Prieto, J. *J. Neurochem.* **57**, 1159–1164 (1991).
31. Nicholls, D. G., Sihra, T. S. & Sánchez-Prieto, J. *J. Neurochem.* **47**, 50–59 (1987).
32. Verhage, M., Besselsen, E., Lopes da Silva, F. H. & Ghijzen, W. E. J. *J. Neurochem.* **53**, 1188–1194 (1989).
33. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. *J. Biol. Chem.* **260**, 3440–3450 (1985).
34. Sims, P. J., Waggoner, A. S., Wang, C. H. & Hoffman, J. F. *Biochemistry* **13**, 3315–3330 (1974).

ACKNOWLEDGEMENTS. We thank E. Lundin for help with manuscript preparation. This work was supported by a grant from the Spanish Ministry of Education and Science. I.H. was supported by a fellowship from Universidad Complutense de Madrid.

Cell-free expression of functional *Shaker* potassium channels

Robert L. Rosenberg & Joan E. East

Departments of Pharmacology and Physiology, CB 7365,
The University of North Carolina at Chapel Hill, Chapel Hill,
North Carolina 27599, USA

THE functional activity of ion channels and other membrane proteins requires that the proteins be correctly assembled in a transmembrane configuration. Thus, the functional expression of ion channels, neurotransmitter receptors and complex membrane-limited signalling mechanisms from complementary DNA has required the injection of messenger RNA or transfection of DNA into *Xenopus* oocytes or other target cells that are capable of processing newly translated protein into the surface membrane^{1–4}. These approaches, combined with voltage-clamp analysis of ion channel currents, have been especially powerful in the identification of structure–function relationships in ion channels^{5–7}. But oocytes express endogenous ion channels^{8,9}, neurotransmitter receptors¹⁰ and receptor–channel subunits¹¹, complicating the interpretation of results in mRNA-injected eggs. Furthermore, it is difficult to control experimentally the membrane lipids and post-translational modifications that underlie the regulation and modulation of ion channels in intact cells. A cell-free system for ion channel expression is ideal for good experimental control of protein expression and modulatory processes. Here we combine cell-free protein translation, microsomal membrane processing^{12–14} of nascent channel proteins, and reconstitution of newly synthesized ion channels into planar lipid bilayers¹⁵ to synthesize, glycosylate, process into membranes, and record *in vitro* the activity of functional *Shaker* potassium channels.

We obtained cDNA for a slowly-inactivating mutant (ShIR; ref. 16) of *Shaker* H4 (ref. 17) that is nearly identical to ShBΔ6-46 (ref. 18). Figure 1 shows the products of *in vitro* translation of ShIR mRNA in the absence and presence of avian oviduct microsomal membranes that are highly efficient at cotranslational processing of transmembrane proteins¹⁹. In the absence of the oviduct microsomes, ShIR mRNA was translated into a roughly 70K (M_r 70,000) protein (lanes 3 and 6), as expected (see legend). When oviduct microsomes were included, a 77K protein was produced in addition to the 70K polypeptide (lanes 2 and 8). The 77K protein predominates in the membrane pellets (lane 8) and was not found in the supernatants (lane 5), indicating that it was localized with the membranes.

The increase in molecular mass from 70K to 77K was due to the glycosylation of the ShIR polypeptide (Fig. 2). Incubation of oviduct microsomes containing the 77K protein (Fig. 2, lane

3) with endoglycosidase H caused the molecular mass to shift to 70K (lane 2). The enzymatically deglycosylated protein was indistinguishable from the 70K unprocessed protein in translation mixtures (lane 4).

To determine whether the products of protein translation were functional *Shaker* K⁺ channels, we incorporated the oviduct microsomes containing the 77K protein into preformed planar lipid bilayers and recorded single-channel currents under voltage-clamp (Fig. 3a). Channel activation required strong depolarizations, with depolarizations up to 0 mV eliciting very little activity (data not shown). As seen previously^{16,18}, channel activation was rapid; in most recordings with channel activity, the first opening event occurred within the 10–20 ms period of

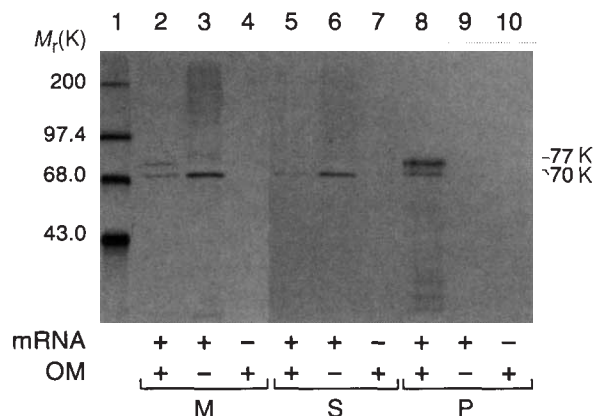


FIG. 1 SDS-polyacrylamide gel analysis (PAGE) of ShIR translation *in vitro*. ShIR is formed by the deletion of amino acids 6–46 of *Shaker* H4 (ref. 17), and is identical to ShBΔ6-46 (refs 18, 27) except for a single amino-acid insertion at position 513 and two amino-acid replacements at positions 583 and 584 (using the original position numbers^{17,27}). Newly synthesized proteins were labelled during translation with [³H]-leucine, separated by SDS-PAGE, and identified by fluorography. Translated proteins in the complete reaction mixture (M), supernatants (S), or pellets (P), in the presence or absence of exogenous mRNA and oviduct microsomes (OM) are shown. In the absence of exogenous mRNA, no proteins were synthesized (lanes 4, 7, and 10), indicating the absence of mRNA in the microsomal preparation or in the reticulocyte lysate. In the presence of ShIR mRNA but absence of oviduct membranes, a polypeptide of M_r 70K was observed (lanes 3 and 6). This value is identical to that calculated by subtracting the molecular mass of residues 6–46 (5K)¹⁷ from that of the primary sequence of *Shaker* H4 (75K)¹⁷. In the presence of ShIR mRNA and oviduct microsomes, a 77K polypeptide is abundant in the membrane pellets (lane 8), and can be seen along with the 70K product in the complete mixture (lane 2). Unlike the 70K polypeptide, the 77K protein is not found in the supernatants (lane 5). The molecular mass is in the range (65–85K) identified in adult *Drosophila* neurons by immunoblots²⁸.

METHODS. Bluescript KS⁺ plasmid containing the inserted cDNA for the ShIR was linearized with *Hind*III and transcription was by the T7 RNA polymerase, using an RNA transcription/capping kit (Stratagene). Avian oviduct microsomes were prepared as described¹⁸ except that Japanese quail rather than chickens were used as the source of oviduct. Oviduct microsomes were treated with staphylococcal nuclease (Boehringer) to degrade endogenous mRNA¹², and were stored at a concentration of 9 mg membrane protein per ml in 250 mM sucrose, 50 mM triethanolamine, 1 mM dithiothreitol, pH 7.5, at –80 °C for up to 18 months. Translation *in vitro* was done with rabbit reticulocyte lysate (Promega) and quail oviduct microsomes. Translation mixtures contained 17.5 μl rabbit reticulocyte lysate, 0.5 nmol mixed amino acids (minus leucine, Promega), 12.5 μCi (88 pmol) [³H]-leucine (NEN), 20 U RNasin (Promega), with or without 1 μg mRNA and 16 μg oviduct membrane protein (as indicated in the figure) in a final volume of 25 μl. Translations were allowed to proceed for 60 min at 30 °C and were halted by placing the reaction mix on ice. After translation, oviduct membranes were pelleted (18,000g for 15 min at 4 °C). Supernatants were removed and the pellets were washed three times in 160 mM NaCl, 10 mM Na-MOPS, pH 7.4. Final supernatants were aspirated, and pellets were dissolved in SDS-PAGE sample buffer. [¹⁴C]-labelled protein standards were also loaded (lane 1). After electrophoresis, gels were treated with Fluoro-Hance (Research Products International), dried, and placed next to Kodak X-Omat film for 24–48 h.

FIG. 2 Carbohydrate added to ShIR translation products by avian microsomes is cleaved by endoglycosidase H (Endo H), an enzyme that cleaves high-mannose oligosaccharides but not complex oligosaccharides from glycoproteins²⁹. Lane 1, ¹⁴C-labelled protein standards. Lanes 2–5 show ³H-labelled ShIR protein obtained in the presence of avian oviduct microsomes. Lane 2, translation products in membrane pellets after treatment with endoglycosidase H, showing the 70K core polypeptide. Lane 3, polypeptides in membrane pellets not treated with endoglycosidase H, showing the predominance of the 77K glycosylated product over the non-glycosylated 70K protein. Lane 4, products in the entire reaction mixture, clearly showing both 77K and 70K proteins. Transmembrane processing and high-mannose glycosylation of the nascent ShIR polypeptide by the oviduct microsomes is evidenced by the increase in molecular mass, the sensitivity of this increase to endoglycosidase H, and the preferential localization of the processed ShIR protein in the membrane pellets.

METHODS. Translations were done as described in Fig. 1. Membrane pellets (lanes 2 and 3) were resuspended in 50 mM sodium acetate, pH 6.0, containing 1% β -mercaptoethanol. SDS was added to a final concentration of 0.1%, providing optimal conditions for the endoglycosidase H (Boehringer). Endoglycosidase H (0.1 U ml⁻¹ final concentration; lane 2) or carrier buffer (50 mM sodium phosphate, pH 7, 0.05% sodium azide; lane 3) was added, and samples were incubated at 37 °C for 2 h. Samples were then treated for SDS-PAGE.

amplifier saturation that followed voltage jumps. In some of the recordings at +20 mV the channel remained open during the entire depolarization, but closed immediately upon repolarization. The absence of detectable 'tail' currents after repolarization, even when the channel was open throughout the depolarization, indicated that channel deactivation was also faster than 10–20 ms, as seen previously^{16,18}. At +40 and +60 mV, channel closings during the test pulse were more frequent, and the channel showed a tendency to enter a long-lasting inactivated state²⁰ before the end of the depolarization (see below). The apparent mean open-time of the channels was >100 ms, a value much greater than the ~4 ms recorded in mRNA-injected oocytes¹⁸. This could be because of the 200 Hz low-pass filtering that made it impossible to resolve brief closing

events (<1 ms), effects of bilayer lipids, incomplete carbohydrate processing, or other differences between oviduct microsomes and *Xenopus* oocytes.

The current-voltage relationship from these experiments is shown in Fig. 3b. The slope of the line indicates a conductance of 14 pS in 50 mM K⁺ (external), 350 mM K⁺ (internal). This is very close to the 13 pS conductance of ShIR K⁺ channels recorded in patches from *Xenopus* oocytes¹⁶. The extrapolated reversal potential is -50 mV, a value identical to the equilibrium potential for K⁺ in these experiments, indicating ideal K⁺-over-Cl⁻ selectivity.

In all 10 blind control experiments, where no mRNA was added to the translation mixture (a fact unknown at the time of bilayer reconstitution), no channel activity of any kind was

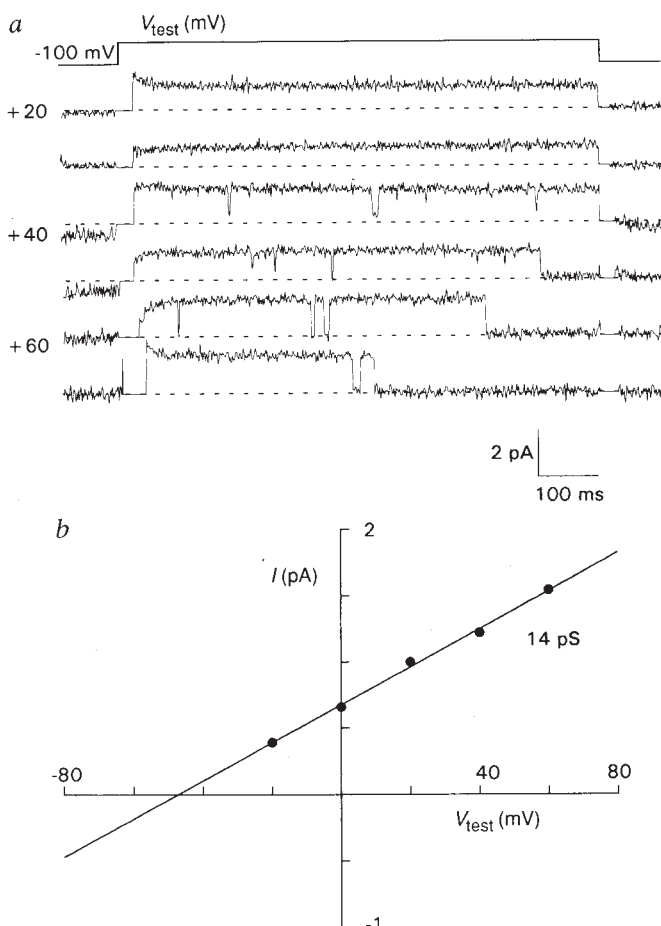


FIG. 3 Single-channel recordings of channels incorporated into planar lipid bilayers from oviduct microsomes containing the 77K protein synthesized from ShIR mRNA. **a**, Selected recordings obtained at the test potentials indicated. The holding potential of -100 mV was maintained for 9.2 s between the 800 ms test pulses. Channel openings are upward; the closed-channel level is indicated by the dashed line. The curvature in the single-channel currents at the beginning of some of the depolarizations is an artifact of the leak subtraction, arising when recordings without channel activity do not perfectly match the time-course of capacitive current decay in the selected recordings. **b**, Current-voltage plot of the single-channel currents; unitary current amplitude is plotted against test potential. The line drawn is the least-squares fit of the data to a straight line. The slope is 14 pS, and the extrapolated reversal potential of -50 mV is exactly the equilibrium potential for K⁺ in these experiments.

METHODS. *In vitro* translation was as described in Fig. 1, with the following exceptions: all volumes were four times greater to increase the amount of protein synthesized; [³H]-leucine was omitted and a mixture of all amino acids except methionine (Promega) was added, providing a complete mix of non-radioactive amino acids; membrane pellets (8,000g, 15 min) were resuspended by gentle trituration in 50 μ l of 160 mM NaCl, 10 mM Na-MOPS, pH 7.4, and allowed to stand at 4 °C for 2 h with occasional gentle trituration to increase the dispersion of the pelleted membranes into the solution. Planar lipid bilayers were formed of 1-palmitoyl-2-oleoyl-phosphatidylethanolamine and 1-palmitoyl-2-oleoyl-phosphatidylserine (Avanti Polar Lipids) (15 mg ml⁻¹ and 5 mg ml⁻¹, respectively, in *n*-decane) across a 150 μ m aperture in a polyvinylidene fluoride partition³⁰. The aqueous solutions contained 350 mM KCl, 1 mM CaCl₂, 10 mM K-HEPES, pH 7.0 (*cis*) and 50 mM KCl, 1 mM CaCl₂, 10 mM K-HEPES, pH 7.0 (*trans*). Oviduct microsomes were added to the *cis* chamber with stirring, and fusion of the microsomes to the bilayers was as described³⁰. Voltage clamp was accomplished with a modified patch-clamp amplifier (Warner Instruments). Currents were filtered at 200 Hz (corner frequency, 8-pole Bessel low-pass) digitized at 1 kHz, and stored in computer memory for off-line analysis. Uncompensated capacitive and resistive currents were subtracted from the recordings shown³⁰. In the configuration of these experiments, the voltages were assigned as *cis* relative to *trans*, and currents flowing from *cis* to *trans* are shown as upward transitions. Current amplitudes were measured by eye with a computer-drawn vector superimposed on the leak-subtracted records. Activation of channels by a membrane depolarization in this configuration suggests that the *cis* chamber represents the cellular interior, consistent with the incorporation of inside-out microsomal membranes by membrane fusion. This is the expected orientation of the microsomes, where the microsomal lumen is topologically equivalent to the cellular exterior.

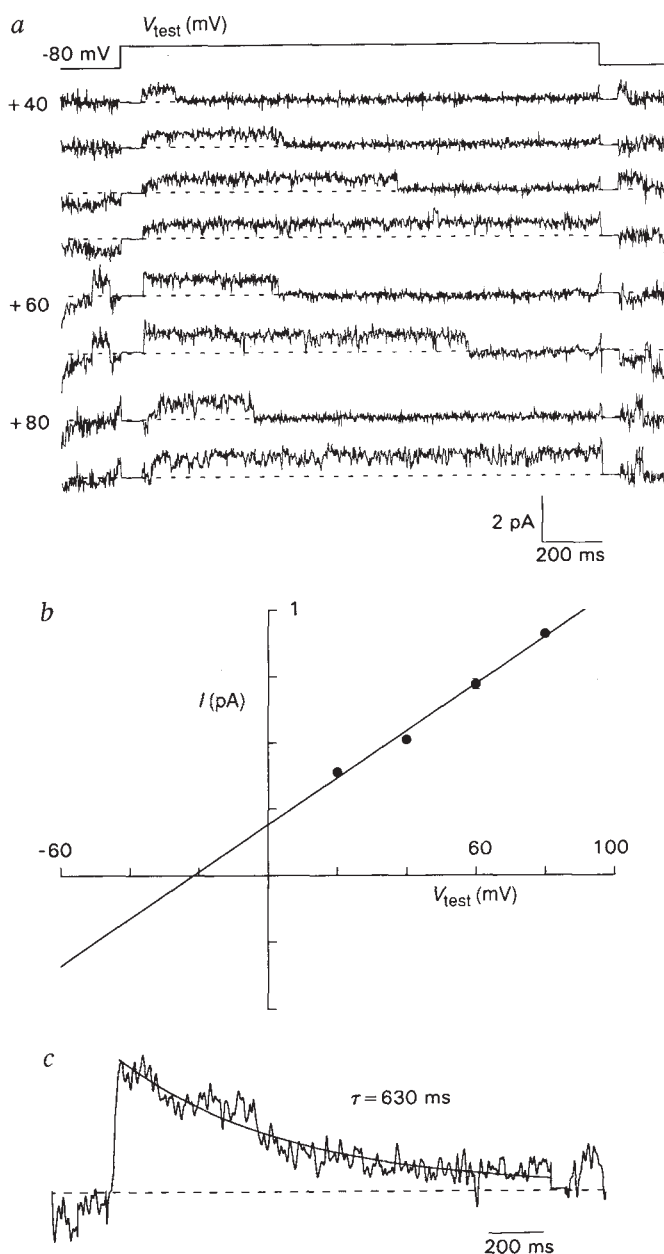


FIG. 4 Analysis of cation selectivity and inactivation kinetics of channels in avian microsomes containing the 77K product. *a*, selected recordings at the test potential indicated. Ionic conditions were 450 mM NaCl, 1 mM CaCl_2 , 10 mM Na-HEPES, pH 7.0, *cis*, and 100 mM KCl, 1 mM CaCl_2 , 10 mM K-HEPES, pH 7.0, *trans*. *b*, Current-voltage plot of the unitary currents against test potential. The conductance (9 pS) is smaller than the 14 pS in Fig. 3, presumably because of the lower K^+ concentration (100 mM compared with 350 mM). The extrapolated reversal potential is -21 mV. Using the equation $E_{\text{rev}} = (25.5 \text{ mV}) \ln((P_{\text{Na}}/P_{\text{K}})([\text{Na}]_o/[\text{K}]_i))$, ref. 26), and substituting 455 mM for $[\text{Na}]_o$ and 105 mM for $[\text{K}]_i$, $P_{\text{K}}/P_{\text{Na}}$ is calculated to be 9.9. *c*, Ensemble averages of channel activity at $+40$ mV. The curve represents the least-squares fit of the data to the equation for a single exponential decay to the zero-current baseline²⁰, yielding a time constant for current inactivation (τ) of 630 msec.

METHODS. Translations were as described in Fig. 3. Planar lipid bilayers were formed and microsomes incorporated as before, except that the holding potential of -80 mV was maintained for 8.4 s between 1.6 s test pulses. Voltage clamp was the same as before, except that in these experiments the voltages were defined as *trans* relative to *cis* and currents from *trans* to *cis* are shown as upward transitions. Thus, the activation of the channels by depolarization suggests that in these experiments the *trans* side represents the cellular interior. This is probably because of the incorporation of microsomes with the reversed topology (inside corresponds to intracellular), a situation that could arise from the mechanical trituration and resuspension of the pelleted microsomes.

recorded. In 6 out of 52 trials with microsomes containing the 77K glycoprotein, there was clear K^+ channel activity, and in an additional 4 experiments some channel activity of approximately 14 pS was recorded, but bilayer noise or breakage made it difficult to assess conductance or selectivity. This provided evidence that the K^+ channel activity was a result of new protein synthesis, and did not derive from a channel endogenous to the avian microsomes. In 1 of the 52 trials, we recorded a channel with very clear Cl^- selectivity, a conductance of 96 pS in 50 mM external, 350 mM internal KCl, and frequent subconductance excursions (30 pS). This channel is almost certainly endogenous to the oviduct microsomes but is present in low abundance.

To determine whether the 14 pS K^+ channels were selective for K^+ over Na^+ , experiments with external NaCl (450 mM) and internal KCl (100 mM) were done (Fig. 4). As before, depolarization of the bilayer evoked channel activity that appeared as outward unitary current events (Fig. 4a). The current-voltage relationship (Fig. 4b) indicates a conductance of 9 pS, and an extrapolated reversal potential of -21 mV. This reversal potential, combined with the lack of Cl^- permeability (Fig. 3), suggests a 10-fold selectivity for K^+ over Na^+ , a value lower than that determined for Shaker A1 channels (selectivity ratio $P_{\text{K}}/P_{\text{Na}} \sim 65$; ref. 21). The measured selectivity ratio is likely to be an underestimate, as the true reversal potential is probably more negative than the extrapolated one because of the expected curvature of the I - V relationship at negative potentials where Na^+ flux is favoured.

The recordings in Fig. 4a show a clear tendency for the channel to enter a long-lasting inactivated state before the end of the long depolarizing test-pulses. This is seen clearly in the ensemble averages of the currents recorded at $+40$ mV (Fig. 4c). The time constant for inactivation is roughly 600 ms, a value similar to those seen when ShBΔ6-46 channels were expressed in frog oocytes (median value $\sim 1,000$ ms)²⁰.

In summary, our results show that functional ShIR K^+ channels can be synthesized, processed, and studied in a completely cell-free system. The molecular mass (77K) is in the range expected, and the 14 pS conductance and the 600 ms inactivation rate were very close to those recorded in *Xenopus* oocytes injected with ShIR mRNA. Channel activation required large depolarizations, and activation and deactivation kinetics were faster than 10–20 ms, as seen previously. Ion selectivity ($P_{\text{K}}/P_{\text{Na}} > 10$) and tetraethylammonium ion-sensitivity (K_i (internal) ~ 1 –2 mM; data not shown) appear similar to those parameters of Shaker K^+ channels in *Xenopus* oocytes.

K^+ channel proteins are probably formed as a tetramer of subunits²², each with six putative membrane-spanning α -helices and one putative membrane-penetrating β -turn^{6,7}. Thus, in addition to being able to cleave signal sequences and transfer the high-mannose carbohydrate from dolichol-polysaccharide to the nascent membrane proteins²³, the oviduct microsomes can promote the trans-membrane folding¹⁴ and oligomerization of complicated membrane proteins. Although carbohydrate may be required for K^+ channel function²⁴, our results indicate that the simple high-mannose carbohydrate is sufficient for most channel behaviour. Post-translational modifications, including protein phosphorylation and carbohydrate processing may also be controlled experimentally, as enzymes, substrates, Golgi membranes²⁵ and activators or inhibitors of the enzymes present in the reticulocyte lysate or oviduct microsomes can be added to translation mixtures. Trials with more complex membrane proteins such as voltage-gated Na^+ and Ca^{2+} channels, where control of intracellular messengers and enzymes is essential for the characterization of channel regulation²⁶, are now possible. □

Received 25 June; accepted 15 September 1992.

1. Lester, H. A. *Science* **241**, 1057–1063 (1988).
2. Perez-Reyes, E. *et al.* *Nature* **340**, 233–236 (1989).
3. Klaiber, K. *et al.* *Neuron* **5**, 221–226 (1990).

4. Yang, X.-C. *et al.* *J. Neurosci.* **12**, 268–277 (1992).
5. Numa, S. *Harvey Lect.* **83**, 121–165 (1989).
6. Guy, H. R. & Conti, F. *Trends Neuro.* **13**, 201–206 (1990).
7. Miller, C. *Science* **252**, 1092–1096 (1991).
8. Dascal, N. *CRC Crit. Rev. Biochem.* **22**, 318–387 (1987).
9. Bourinot, E., Fournier, F., Nargeot, J. & Charpent, P. *FEBS Lett.* **299**, 5–9 (1992).
10. Fuharty, S. J., Reagan, L. P. & White, M. M. *J. Neurochem.* **56**, 1307–1311 (1991).
11. Buller, A. L. & White, M. M. *Molec. Pharmac.* **37**, 423–428 (1990).
12. Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 84–93 (1983).
13. Walter, P., Gilmore, R. & Blobel, G. *Cell* **38**, 5–8 (1984).
14. Kobilka, B. K. *J. Biol. Chem.* **265**, 7610–7618 (1990).
15. Miller, C. *Ion Channel Reconstitution* (Plenum, New York, 1986).
16. Yellen, G., Jurman, M. E., Abramson, T. & MacKinnon, R. *Science* **251**, 939–942 (1991).
17. Kamb, A., Tseng-Crank, J. & Tanouye, M. A. *Neuron* **1**, 421–430 (1988).
18. Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Science* **250**, 533–538 (1990).
19. Das, R. C., Brinkley, S. A. & Heath, E. C. *J. Biol. Chem.* **255**, 7933–7940 (1980).
20. Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Neuron* **7**, 547–556 (1991).
21. Iverson, L. E., Tanouye, M. A., Lester, H. A., Davidson, M. & Rudy, B. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5723–5727 (1988).
22. MacKinnon, R. *Nature* **350**, 232–235 (1991).
23. Alberts, B. *et al.* *Molecular Biology of the Cell* 2nd edn (Garland, New York, 1991).
24. Zona, C., Eusebi, F. & Miledi, R. *Proc. R. Soc. B* **239**, 119–127 (1990).
25. Segev, N. *Science* **252**, 1553–1556 (1991).
26. Hille, B. *Ionic Channels of Excitable Membranes* 2nd edn (Sinauer, Sunderland, Massachusetts, 1992).
27. Schwartz, T. L., Tempel, B. L., Papazian, D. M., Jan, Y. N. & Jan, L. Y. *Nature* **331**, 137–142 (1988).
28. Schwartz, T. L., Papazian, D. M., Carretto, R. C., Jan, Y.-N. & Jan, L. Y. *Neuron* **2**, 119–127 (1990).
29. Maley, F., Trimble, R. B., Tarentino, A. L. & Plummer, T. H. Jr *Analyt. Biochem.* **180**, 195–204 (1989).
30. Rosenberg, R. L. & Chen, X.-h. *J. gen. Physiol.* **97**, 1207–1225 (1991).

ACKNOWLEDGEMENTS. We thank T. Slopes for the quail oviduct, R. MacKinnon for SHIR cDNA, E. Azhderian for suggesting the idea and R. Nicholas, D. Lamson, P. Koplas and J. Haack for discussion. This work was supported by the NIH and the American Heart Association. R.L.R. is an Established Investigator of the American Heart Association.

Function of a truncated dihydropyridine receptor as both voltage sensor and calcium channel

Kurt G. Beam*, Brett A. Adams*, Tetsuhiro Niidome†, Shosaku Numa†‡ & Tsutomu Tanabe†§

*Department of Physiology, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, Colorado 80523, USA

†Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Kyoto 606, Japan

THE skeletal muscle dihydropyridine (DHP) receptor serves dual functions, as a voltage sensor for excitation–contraction coupling and as an L-type calcium channel^{1–3}. Biochemical analysis indicates the presence of two forms of the DHP receptor polypeptide in skeletal muscle, a full-length translation product present as a minor species and a much more abundant form that has a truncated carboxy-terminus^{4–6}. On the basis of these and other observations⁷, it has been proposed⁸ that, in skeletal muscle, only the full-length DHP receptor can function as a calcium channel and that the truncated form can only function as a voltage sensor for excitation–contraction coupling. To resolve this issue, we have now constructed a complementary DNA (pC6Δ1) encoding a protein corresponding to the truncated DHP receptor in skeletal muscle. Expression of pC6Δ1 in dysgenic myotubes fully restores both excitation–contraction coupling and calcium current, consistent with the idea that a single class of DHP receptors performs both functions.

The expression plasmid pCAC6 encodes a full-length DHP receptor of 1,873 amino acids². The newly constructed pC6Δ1 encodes a DHP receptor of 1,662 amino acids; this protein lacks the final 211 C-terminal amino acids of the protein encoded by pCAC6, but is otherwise identical (Fig. 1). The amino-acid sequence TNPLAR (single-letter code; residues 1,653–1,658) is detected in tryptic fragments of the predominant form of the

DHP receptor present in skeletal muscle². Moreover, site-specific antibodies indicate that the predominant form is truncated between residues 1,685 and 1,699 (ref. 6). Thus, the polypeptide encoded by pC6Δ1 seems to correspond closely to the predominant form of the DHP receptor, although it is perhaps somewhat shorter.

Dysgenic myotubes (from muscle precursor cells) were injected with pC6Δ1 and tested 2–3 days later for restored excitation–contraction (EC) coupling³, as indicated by contraction in response to electrical stimulation. Of the ~150 myotubes that survived injection with pC6Δ1, 53 were found to have restored EC coupling, a fraction comparable to that observed for dysgenic myotubes injected with pCAC6 (ref. 9). Figure 2a illustrates electrically evoked contractions of a dysgenic myotube expressing pC6Δ1. Contractions of this myotube could be elicited in the absence of external calcium or in the presence of 0.5 mM external cadmium, indicating that the truncated DHP receptor encoded by pC6Δ1 mediates skeletal muscle-type EC coupling, which does not require the entry of external calcium¹⁰. Results similar to those shown in Fig 2a were obtained in every pC6Δ1-injected dysgenic myotube that was examined (*n* = 26).

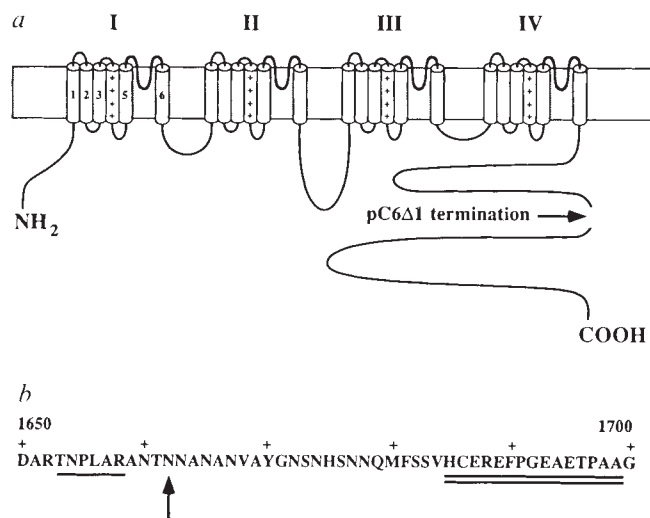


FIG. 1 a, Schematic representation of the full-length (1,873 amino acids) skeletal muscle DHP receptor² and the mutant DHP receptor in which the C terminus has been truncated (arrow). The four repeating units of homology are labelled I–IV, the predicted transmembrane segments (S1–S3, S5, S6) are labelled for the first repeat, and the S4 segment, postulated to be the voltage sensing region², is indicated by plus signs in all four repeats. b, The amino-acid sequence (single-letter code) in the region of truncation (numbers refer to the full-length protein encoded by the plasmid pCAC6). The arrow indicates the last residue (asparagine 1,662) encoded by pC6Δ1. From residue 1,650 to 1,700, every tenth residue is indicated by a plus sign. The peptide fragment TNPLAR (underlined) has been detected by direct sequencing of tryptic fragments² of the purified skeletal muscle DHP receptor. The doubly underlined sequence indicates the region of termination of the predominant form of DHP receptor found in skeletal muscle, as determined by site-directed antibodies⁶.

METHODS. The expression plasmid pC6Δ1, carrying the cDNA encoding a truncated form of the rabbit skeletal muscle DHP receptor, was prepared as follows: the 1.7-kilobase pair (kb) *Pst*II(3,998)–*Hind*III (3'-terminal linker) fragment from pCAC6 (ref. 2) was ligated with the *Pst*II–*Hind*III-cleaved vector M13mp18. A 32-base mutagenic oligonucleotide having the sequence 5'-GACACCACAGCCATCAGTTGGATTGGCAGCA-3' was synthesized with an automatic DNA synthesizer (Applied Biosystems); this oligonucleotide corresponds to nucleotide residues 4,971–4,986 and 5,620–5,635 of pCAC6. Oligonucleotide-directed *in vitro* mutagenesis was done to produce a deletion mutant¹⁵. The resulting 233-base pair (bp) *Aat*I–*Hind*III fragment, having a 633-bp deletion, was ligated with the 4.5-kb *Hind*III (5'-terminal linker)–*Bgl*II(4,488) fragment and the *Bgl*II(4,488)–*Aat*I(4,809) fragment from pCAC6 and *Hind*III-cleaved pKCRH2 (ref. 16) to yield the plasmid pC6Δ1. To ensure that no other mutation occurred during mutagenesis, we sequenced the region extending from the *Aat*I(4,809) site to the *Hind*III (3'-terminal linker) site of pC6Δ1.

‡ Deceased.

§ Present address: Howard Hughes Medical Institute and Department of Cellular and Molecular Physiology, Boyer Center for Molecular Medicine, Yale University School of Medicine, PO Box 9812, New Haven, Connecticut 06536-0812, USA.

Expression of pC6Δ1 restored the intramembrane charge movement¹¹ (Fig. 2b) that has been implicated as representing voltage-driven conformational changes of the DHP receptor functioning as the voltage sensor for EC coupling^{1,12}. The maximum immobilization-resistant charge movement ($Q_{\max}$) in 10 dysgenic myotubes expressing pC6Δ1 had a mean ($\pm$ s.d.) of $5.9 \pm 2.0 \text{ nC } \mu\text{F}^{-1}$, close to the value of $6.4 \pm 2.8 \text{ nC } \mu\text{F}^{-1}$ ($n = 10$) for $Q_{\max}$ in dysgenic myotubes expressing pCAC6 (ref. 12). The charge movement restored by pC6Δ1 was sensitive to the DHP antagonist (+)-PN 200-110 (Fig. 2b). In four myotubes expressing pC6Δ1, $10 \mu\text{M}$ (+)-PN 200-110 reduced $Q_{\max}$ from 6.5 ± 3.0 to $4.7 \pm 2.4 \text{ nC } \mu\text{F}^{-1}$, a reduction of $1.7 \pm 0.6 \text{ nC } \mu\text{F}^{-1}$. The fractional reduction (28%) is very similar to that produced in total charge movement in adult mammalian skeletal muscle by application of DHP antagonists¹³. In 11 dysgenic myotubes expressing pCAC6, $10 \mu\text{M}$ (+)-PN 200-110 reduced $Q_{\max}$ from 6.2 ± 2.2 to $4.7 \pm 1.6 \text{ nC } \mu\text{F}^{-1}$. Thus, the quantity and DHP-sensitivity of immobilization-resistant charge movement were similar in myotubes expressing pC6Δ1 and pCAC6. Taken together, our results indicate that the truncated DHP receptor is able to function as a voltage sensor for skeletal muscle-type EC coupling.

The truncated DHP receptor is also able to function as a calcium channel, as demonstrated by the restoration of a slowly-activating calcium current in dysgenic myotubes injected with pC6Δ1 (Fig. 3a). The calcium current restored by pC6Δ1 had a density of $3.7 \pm 1.8 \text{ pA pF}^{-1}$ ($n = 19$) and a time constant of activation (τ_{act} ; see ref. 14), measured at the test potential ($\approx 40 \text{ mV}$) that elicited maximal inward current, of $68 \pm 38 \text{ ms}$ ($n = 19$). For comparison, the calcium current in dysgenic myotubes expressing pCAC6 has been reported to have a similar density, $4.4 \pm 2.7 \text{ pA pF}^{-1}$ (ref. 9), as well as a similar τ_{act} ,

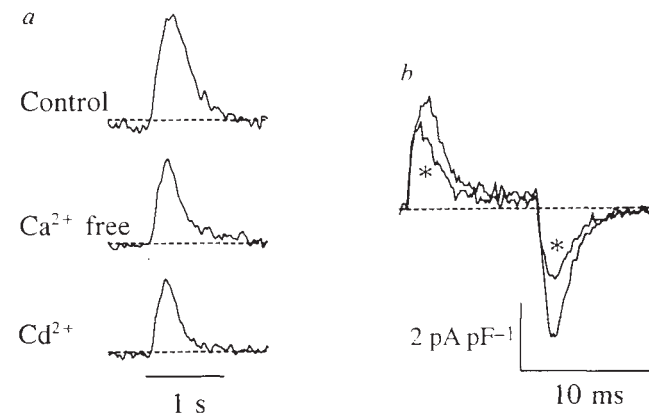


FIG. 2 Restoration of electrically evoked contractions (a) and immobilization-resistant charge movement (b) in dysgenic myotubes expressing pC6Δ1. a, Contractions were initially recorded in normal rodent saline (Control), then in Ca^{2+} -free saline (Ca^{2+} free, made by equimolar substitution of Mg^{2+} for Ca^{2+}), and finally in 0.5 mM Cd^{2+} -containing saline (Cd^{2+}). Vertical scale, arbitrary units. b, Immobilization-resistant charge movements were measured before (unlabelled trace) and during (*) exposure to $10 \mu\text{M}$ (+)-PN 200-110. Dysgenic myotube BP81, linear capacitance (C) = 510 pF , test potential of $+40 \text{ mV}$.

METHODS. The procedures for extracellular stimulation and optical recording of contractions were as described previously³. The rodent saline contained (mM): 146 NaCl, 5 KCl, 2 CaCl_2 , 1 MgCl_2 and 10 HEPES buffer, pH 7.4 with NaOH. Whole-cell¹⁷ ionic and capacitive currents were measured at 20 – 22°C as described previously^{3,12}. Steady holding potential was -80 mV . Patch pipettes contained (in mM) 140 caesium aspartate, 10 HEPES, 10 Cs_2EGTA , 5 Mg_2Cl , pH 7.4 with CsOH. The bath contained (in mM) 145 tetraethylammonium chloride, 10 HEPES, 10 CaCl_2 , 0.003 tetrodotoxin, pH 7.4. For the measurement of calcium currents, test pulses were preceded by a 1 s step to -30 mV to inactivate endogenous T-type calcium current¹⁸. For the measurement of charge movements, calcium-channel ionic currents were blocked by the addition of 0.5 mM CdCl_2 and 0.1 mM LaCl_3 to the bath solution and the pulse protocol¹² consisted of a 1 s prepulse to -30 mV and subsequent 20 – 30 ms repolarization to -50 mV , followed by depolarization to the desired test potential.

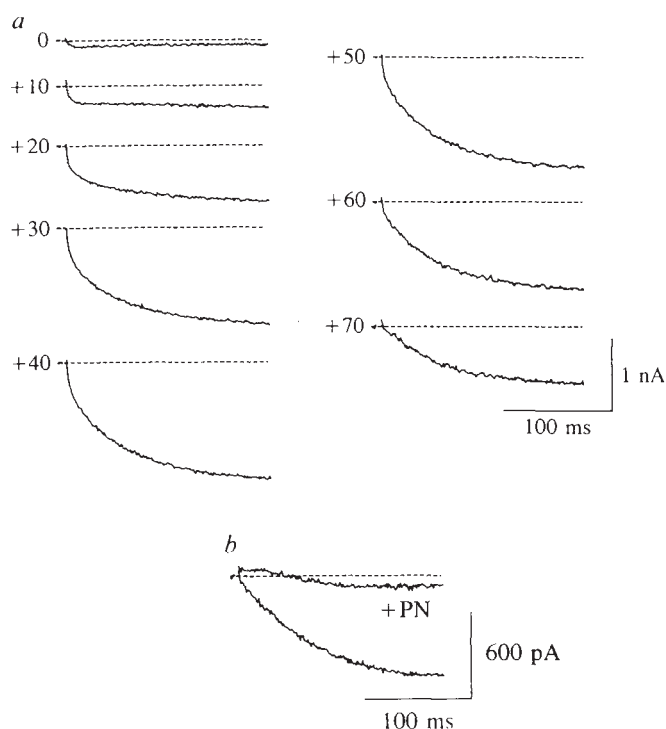


FIG. 3 Calcium currents restored in dysgenic myotubes by expression of pC6Δ1. a, Calcium currents for varying test potentials (mV), as indicated. Cell BL12, $C = 465 \text{ pF}$. b, Calcium currents elicited by a test potential of $+40 \text{ mV}$ before and during exposure to $1 \mu\text{M}$ (+)-PN 200-110. Cell BP80, $C = 280 \text{ pF}$.

$41.3 \pm 27.4 \text{ ms}^{14}$. As previously shown for the slowly-activating calcium current in dysgenic myotubes expressing pCAC6 (ref. 3), the calcium current produced by expression of pC6Δ1 was substantially blocked by $1 \mu\text{M}$ (+)-PN 200-110 (Fig. 3b).

If restored charge movement is taken as an indirect measure of the number of DHP receptors expressed in an injected myotube, then the ratio of maximal L-type calcium conductance ($G_{\max}$) to restored charge movement gives an indication of the ability of the expressed DHP receptor to function as a calcium channel¹². Restored charge movement can be defined as $Q'_{\max} = Q_{\max} - 2.5 \text{ nC } \mu\text{F}^{-1}$, where $2.5 \text{ nC } \mu\text{F}^{-1}$ is the average, maximal, immobilization-resistant charge movement in non-injected dysgenic myotubes¹². In the case of myotubes expressing pC6Δ1, $G_{\max}$ was $136 \pm 84 \text{ nS nF}^{-1}$ ($n = 10$) and $Q'_{\max}$ was $3.4 \text{ nC } \mu\text{F}^{-1}$, yielding a $G_{\max}/Q'_{\max}$ ratio of 40 nS pC^{-1} . For comparison, $G_{\max}/Q'_{\max}$ for myotubes expressing pCAC6 was 36 nS pC^{-1} (ref. 12). Thus, DHP receptors produced by expression of the full-length or truncated cDNAs seem to differ little in their ability to function as calcium channels.

Our results demonstrate that a truncated DHP receptor, which probably corresponds closely in length to the predominant form detected biochemically in skeletal muscle, is able to function not only as a voltage sensor for EC coupling but also as a calcium channel. This truncated DHP receptor encoded by pC6Δ1 lacks three consensus sites for phosphorylation by cyclic AMP-dependent protein kinase² (residues 1,757, 1,772 and 1,854) which are present in the full-length form encoded by pCAC6. Clearly, a DHP receptor lacking these three sites can carry out the essential functions of both voltage sensor and calcium channel. Although we cannot conclude that individual DHP receptors function simultaneously as channel and voltage sensor, our results are consistent with the hypothesis that both of these functions are performed in skeletal muscle by a single molecular species of the DHP receptor. This hypothesis accounts well for all the similarities observed in pCAC6- and pC6Δ1-transfected dysgenic myotubes. □

Received 16 June; accepted 21 September 1992.

1. Rios, E. & Brum, G. *Nature* **325**, 717–720 (1987).
2. Tanabe, T. *et al.* *Nature* **328**, 313–318 (1987).
3. Tanabe, T., Beam, K. G., Powell, J. A. & Numa, S. *Nature* **336**, 134–139 (1988).
4. De Jongh, K. S., Merrick, D. K. & Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8585–8589 (1989).
5. Lai, Y., Seagar, M. J., Takahashi, M. & Catterall, W. A. *J. biol. Chem.* **265**, 20839–20848 (1990).
6. De Jongh, K. S., Warner, C., Colvin, A. A. & Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10778–10782 (1991).
7. Schwartz, L. M., McCleskey, E. W. & Almers, W. *Nature* **314**, 747–751 (1985).
8. Catterall, W. A. *Cell* **64**, 871–874 (1991).
9. Tanabe, T., Mikami, A., Numa, S. & Beam, K. G. *Nature* **344**, 451–453 (1990).
10. Armstrong, C. M., Bezanilla, F. M. & Horowitz, P. *Biochim. biophys. Acta* **267**, 605–608 (1972).
11. Schneider, M. F. & Chandler, W. K. *Nature* **242**, 244–246 (1973).
12. Adams, B. A., Tanabe, T., Mikami, A., Numa, S. & Beam, K. G. *Nature* **346**, 569–572 (1990).
13. Lamb, G. D. & Walsh, T. *J. Physiol., Lond.* **393**, 595–617 (1987).
14. Tanabe, T., Adams, B. A., Numa, S. & Beam, K. G. *Nature* **352**, 800–803 (1991).
15. Taylor, J. W., Otto, J. & Eckstein, F. *Nucleic Acids Res.* **13**, 8764–8785 (1985).
16. Mishina, M. *et al.* *Nature* **307**, 604–608 (1984).
17. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
18. Nowicky, M. C., Fox, A. P. & Tsien, R. W. *Nature* **316**, 440–443 (1985).

ACKNOWLEDGEMENTS. We thank L. Bennett and A. Beam for help in providing cultured myotubes and J. Nakai for performing some of the injections of pC6Δ1. This research was supported in part by the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation, the Japanese Foundation of Metabolism and Diseases and by the NIH. B.A.A. was supported by a postdoctoral fellowship from the American Heart Association (Colorado Affiliate).

Proteasome subunits encoded in the MHC are not generally required for the processing of peptides bound by MHC class I molecules

Daniele Arnold*, James Driscoll†, Matthew Androlewicz‡, Eric Hughes‡, Peter Cresswell‡ & Thomas Spies*

* Division of Tumor Virology, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA

† Department of Cellular and Molecular Physiology, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Howard Hughes Medical Institute, Section of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

ANTIGEN processing provides major histocompatibility complex (MHC) class I molecules with short peptides, which they selectively bind and present to cytotoxic T lymphocytes^{1–4}. The proteolytic system generating these peptides in the cytosol is unidentified, but their delivery into the endoplasmic reticulum is mediated by the TAP1–TAP2 transporter encoded in the MHC class II region^{5–9}. Closely linked to *TAP1* and *TAP2* are genes for the *LMP2* and *LMP7* proteins¹⁰, which resemble components of proteasomes^{11–13}, proteolytic complexes known to degrade cytosolic proteins¹⁴. This association has led to the common assumption that proteasomes function in this immunological pathway (discussed in ref. 15). We now show that the expression of stably assembled class I molecules and apparently normal peptide processing can be completely restored in the absence of *LMP2* and *LMP7* in the human lymphoblastoid cell line mutant 721.174 (refs 16, 17). The identity of *LMP7* is directly confirmed by reconstitution of a proteasomal subunit after gene transfer. These results therefore dispute the hypothetical involvement of proteasomes in antigen processing, although a more subtle effect of *LMP2* and *LMP7* cannot be ruled out.

In several lymphoblastoid cell line (LCL) mutants, defects in either of the *TAP1* or *TAP2* genes impair the assembly and surface expression of MHC class I molecules and their ability to present intracellular antigens to cytotoxic T lymphocytes^{5–9}. This phenotype results from a lack of cytosolic peptides of appropriate length in the endoplasmic reticulum, where they become stably associated with class I molecules^{2,5–9,18–20}. Similar cell line mutants with lesions in *LMP2* or *LMP7* have not been identified, but all of these genes are deleted from both MHC

haplotypes in the human LCL mutant 721.174 (refs. 16, 17). We therefore investigated the function of *LMP2* and *LMP7* by sequential gene transfer into .174 cells. In these experiments, human complementary DNA clones under the transcriptional control of the Rous sarcoma virus (RSV) promoter were transferred into .174 cells and stably transfected cell populations were selected using dominant drug-resistance markers.

At least two subunits of 20S proteasomes are missing in the .174-derived T2 cells²¹ and genetic and biochemical data indicate that these correspond to the *LMP2* and *LMP7* gene products^{11–13,21,22}. But this has not been directly demonstrated by reconstitution of the incomplete proteasomes. We therefore transfected .174 cells with an RSV.5(neo)–*LMP7* plasmid construct²³. The *LMP7*–E2 complementary DNA of 1,077 base-pairs (bp) corresponds to RING10 (ref. 12), but it includes a distinct 5'-end translated sequence resulting from alternative exon usage²². A transfectant isolate expressing *LMP7* messenger RNA was compared to the LCL wild-type 721.45 (ref. 16) and mutant .174 cells by two-dimensional gel electrophoresis of immunoprecipitated 20S proteasomes. The results showed the precise restoration of a distinct subunit previously specified as *LMP7* (refs 11, 21, 22), which is therefore encoded by the *LMP7* gene (Fig. 1).

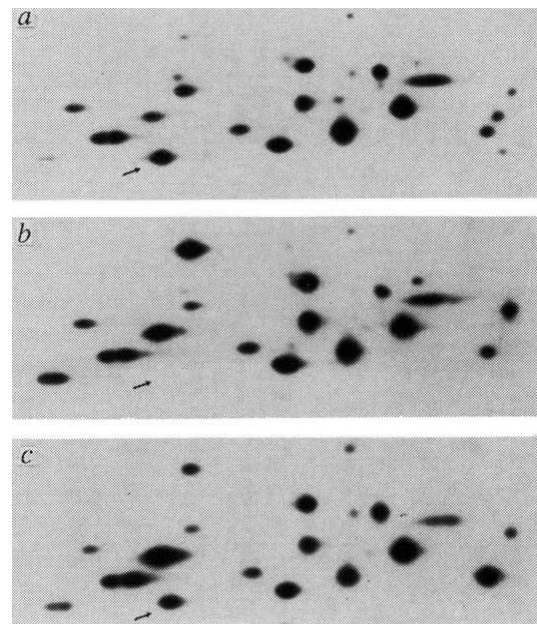


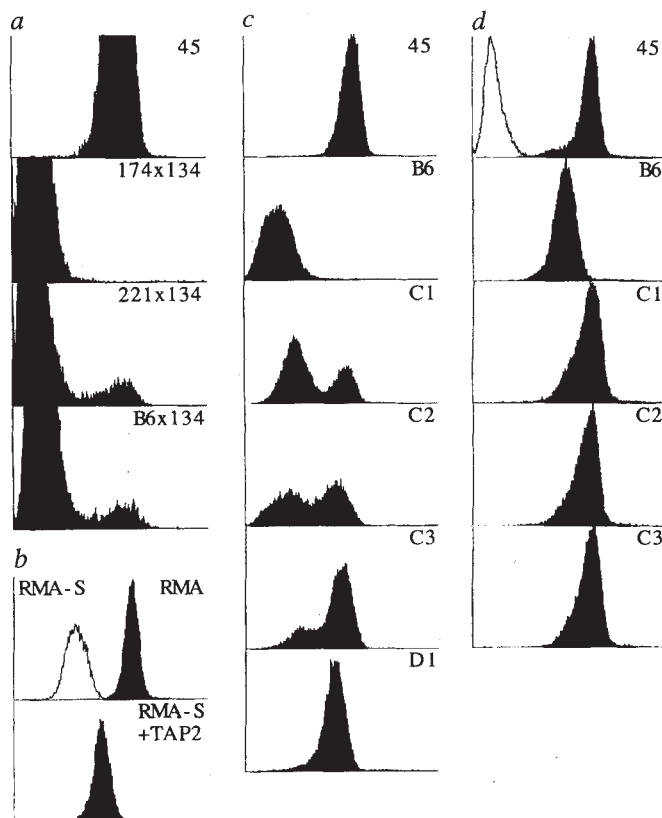
FIG. 1 Reconstitution of the *LMP7* proteasomal subunit in mutant .174 cells by transfection and expression of the *LMP7*–E2 cDNA. The subunit profiles of metabolically labelled 20S proteasomes were generated by immunoprecipitation with a specific antiserum followed by two-dimensional gel electrophoresis. Arrows on fluorographs identify the *LMP7* polypeptide, which is present in .45 cells (a), missing in .174 cells²¹ (b), and restored in the .174-derived transfectant cells (c).

METHODS. The anti-proteasome antiserum was produced by immunizing a New Zealand White rabbit with emulsified 20S proteasome from human erythrocytes, which was purified by repeated gel filtration as described³³. Enzyme purity was assessed by non-denaturing polyacrylamide gel electrophoresis (PAGE) and coincident protease activity³³. SDS-PAGE of these preparations gave the characteristic set of subunits with *M_s*s ranging from 20 to 32 kD³³. For immunoprecipitation, 10⁷ cells were metabolically labelled in the presence of [³⁵S]methionine (500 μCi) in 2 ml RPMI (minus methionine) medium for 4 h and further incubated for 2 h after addition of 6 ml complete culture medium. Detergent lysates³⁰ were pre-adsorbed with normal rabbit serum (20 μl) and insoluble protein A (100 μl; Sigma) for 16 h. Specific antiserum was added to cleared lysates at 10 μg ml⁻¹ for 90 min. Immunocomplexes were then isolated with protein A–Sepharose, washed repeatedly and analysed by non-equilibrium pH-gradient gel electrophoresis (using pH 3.5–10 Ampholines; Pharmacia) in the first dimension, followed by SDS-PAGE (12%) in the second dimension³⁴. Fixed gels were treated with Amplify (Amersham), dried and exposed for 6 days to Kodak XAR film.

FIG. 2 Gene transfer of TAP1 and TAP2 restores normal surface levels of class I molecules in mutant .174 cells. Cloned human cDNAs in plasmid vectors without episomal replication sequences were transferred into .174 cells and stably transfected cell populations were examined for HLA-B5 (a, c) and HLA-A2 (d) surface expression by flow cytometry. a, TAP1 is functional in the .174-derived B6 cells, by transcomplementation of mutant .134 cells: B6 $\times$.134 heterokaryons generated by cell fusion bound the 4D12 monoclonal antibody (anti-HLA-B5) (ref. 26) in equal amounts as wild-type .45 cells. The frequency of positive events (4.6%) was similar to the control fusion of .221 $\times$.134 cells (4.8%). Cell fusions of .174 $\times$.134 and B6 $\times$ B6 (not shown) gave negative results. b, Transfection of RSV.5(neo)-TAP2 into mouse mutant RMA-S cells resulted in increased binding of the 28.14.8S monoclonal antibody (anti-H-2D^b)³⁵, at 25–30% of the level on wild-type RMA cells. c, The C1–C3 cell populations isolated by transfection of B6 cells with RSV.5(neo)-TAP2 showed fluorescence intensities similar to .45 cells after staining with the 4D12 monoclonal antibody; the D1 isolate was obtained by transfection with pREP8Δ-TAP2. d, Transfectants C1–C3, purified by limiting dilution cloning, bound the MA2.1 monoclonal antibody (anti-HLA-A2)²⁷ in amounts comparable to .45 cells.

METHODS. RSV.5(gpt)-TAP1 and B6 cells have been described⁶. TAP2 cDNA²⁵ was ligated into RSV.5(neo)²³ and pREP8Δ using flanking *Xba*I and *Xho*I restriction sites, respectively; pREP8Δ was modified from pREP8 (Invitrogen) by deletion of a 4,213-bp *Bst*Ell-*Eco*RV restriction fragment containing the Epstein-Barr virus origin of replication and nuclear antigen-1 (*EBNA-1*) gene. Cell fusions were prepared with polyethylene glycol (Boehringer) according to standard protocols³⁴ and examined by indirect immunofluorescence and flow cytometry (see below) 18 h after fusion. Non-viable cells were excluded by propidium iodide staining³⁴. Transfections were carried out with a Gene Pulser (Bio-Rad) in 0.4-cm electrode cuvettes at 210 V and 960 μF (250 V and 500 μF with RMA-S cells) as described⁶. Transfectants were selected in 24-well plates, with G418 (Gibco) (450 μg ml⁻¹; 1 mg ml⁻¹ with RMA-S cells) for RSV.5(neo) and with L-histidinol (Sigma) (2.5 mM) for pREP8Δ. After 4–6 weeks of selection, samples of drug-resistant cell populations were stained by indirect immunofluorescence³⁴ with FITC-conjugated goat anti-mouse F(ab')₂ (Cappel) as the second layer and examined with a Becton Dickinson FACSscan flow cytometer.

By reconstituting TAP1 and TAP2, we aimed at isolating an LMP2 and LMP7 null phenotype. The TAP1 subunit is synthesized in normal amounts in the .174-derived B6 cells transfected with the RSV.5(gpt)-TAP1 plasmid construct⁹. As in mutant .174 cells, HLA-B5 is absent from the surface of B6 cells and HLA-A2 is reduced to about 40% of the wild-type level on .45 cells⁶ (Fig. 2c, d). In a functional assay, by transcomplementation of LCL mutant 721.134 cells (TAP1⁻) (refs 6, 9) after cell fusion, B6 cells were equally effective as the control LCL 721.221 cells (TAP1⁺, MHC class I null)²⁴ (Fig. 2a). A control transfection of the full-length TAP2 cDNA²⁵ in RSV.5(neo) into mouse mutant RMA-S cells (TAP2⁻) (refs 5, 7) gave increased surface levels of H-2D^b in 22 of 36 drug(G418)-resistant isolates (Fig. 2b). Following transfection of B6 cells with RSV.5(neo)-TAP2, we obtained 30 independent G418-resistant cell populations. One of these unexpectedly included a 5% proportion of cells with surface HLA-B5, as revealed by binding of the 4D12 monoclonal antibody²⁶ and flow cytometry. Subsequent transfections again yielded HLA-B5⁺ cells (10–80%) in 5 out of 107 polyclonal isolates; antibody-staining profiles of the C1–C3

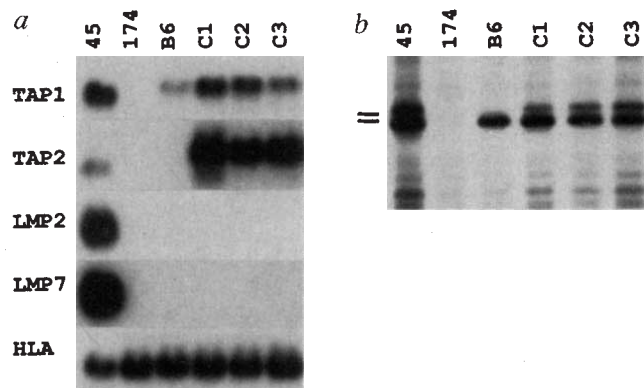


isolates are shown in Fig. 2c. The unusually low frequency of these cells and their very slow proliferation were associated with sequences in the RSV.5(neo) vector, as similar HLA-B5⁺ cells with normal viability occurred in 20 of 36 histidinol-resistant B6 cell populations transfected with pREP8Δ-TAP2 under identical conditions (Fig. 2c). Contrasting these results, no HLA-B5⁺ cells were detected after transfection of mutant .174 cells with RSV.5(neo)-TAP2. The HLA-B5⁺ phenotype was thus reproducibly generated by co-expression of TAP1 and TAP2. This was confirmed by the presence of vector-derived TAP1 and TAP2 mRNA in clonal isolates of the transfectant C1–C3 cells (Fig. 3a) and by immunoprecipitation of TAP1–TAP2 polypeptide complexes (Fig. 3b).

The levels of HLA-B5 on the surface of the transfectant C1–C3 cells and of all of the other positive isolates were about as high as on wild-type .45 cells, ranging from 60 to 100% under standard culture conditions at 37 °C (Fig. 2c). Correspondingly, HLA-A2 and autologous β_2 -microglobulin (β_2m) were increased to normal amounts, as measured by binding of the specific MA2.1 and BBM1 monoclonal antibodies^{27,28}, respectively (Fig. 2d;

FIG. 3 a, Detection of vector-derived TAP1 and TAP2 mRNA in the C1–C3 double transfectants after cell cloning. The TAP1 and TAP2 transcripts were 120 bp longer than the corresponding mRNAs in .45 cells, because of vector 3'-end maturation sequences. No bands were detected by hybridization with LMP2 and LMP7 cDNA probes. All lanes contained similar amounts of control HLA-B mRNA. Differences in intensities between the parallel experiments are not representative. b, Immunoprecipitation of TAP1–TAP2 complexes from various cell lysates using a specific anti-TAP1 antiserum. The positions of the TAP1 and TAP2 polypeptides²⁹ are indicated by the lower and the upper bar, respectively.

METHODS. Preparation and blot hybridization of total RNA (25 μg per lane) and labelling of cDNA probes with [³²P]dCTP by random oligonucleotide priming were according to standard procedures. For protein analysis, detergent lysates of 10⁷ cells metabolically labelled with [³⁵S]methionine (250 μCi for 4 h) were prepared and precleared as described³⁰. Affinity-purified anti-TAP1 antiserum⁹ was used at 10 μg ml⁻¹. Isolated immunocomplexes were separated by SDS-PAGE (10%).



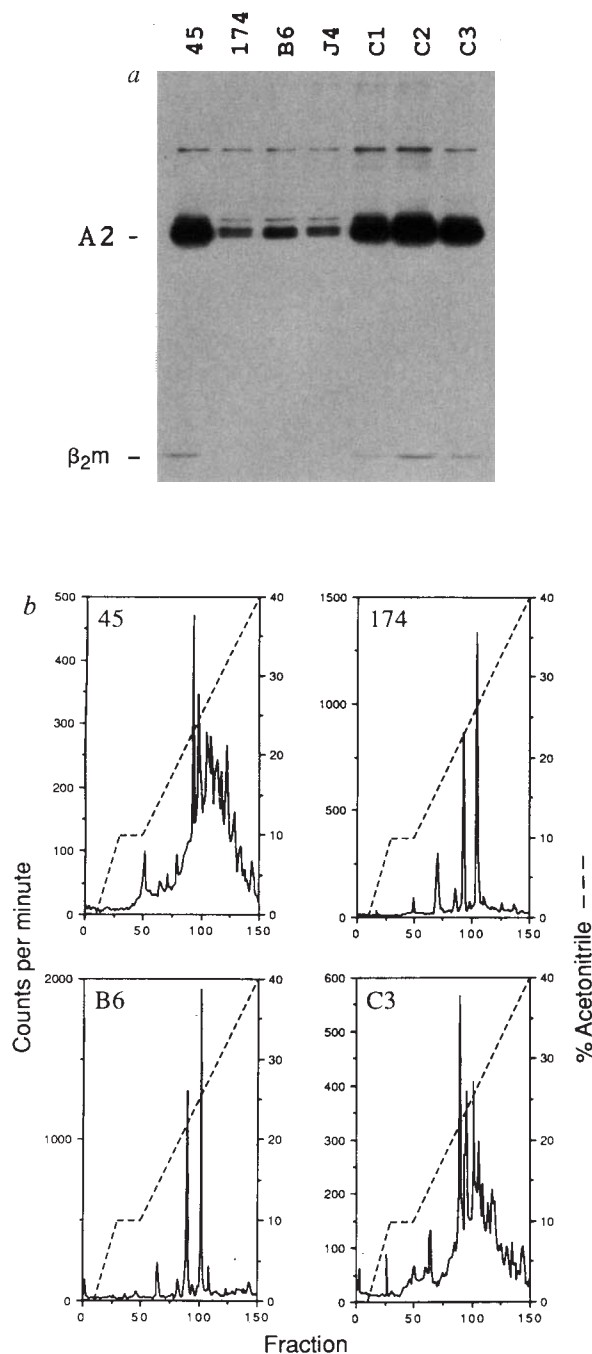


FIG. 4 Stable assembly and normal peptide loading of HLA-A2 molecules in the absence of LMP2 and LMP7. *a*, HLA-A2 heavy chain-β₂m complexes in cell lysates of the C1-C3 cloned transfectants were as stable as from .45 cells after overnight dissociation at 4 °C. Much less material was recovered from lysates of .174 cells and the derivative B6 cells (TAP1⁺) and J4 cells (TAP2⁺); J4 cells were isolated by transfection with RSV.5(neo)-TAP2 and selection for expression of TAP2 mRNA. A2, HLA-A2 heavy chain. *b*, HPLC profiles of peptides bound to HLA-A2 from .45 and mutant .174 cells and the transfectant B6 and C3 cells.

METHODS. Detergent lysates of samples of 10⁷ cells labelled with [³⁵S]methionine (100 μCi for 90 min) were prepared and processed as described³⁰. Purified MA2.1 monoclonal antibody²⁷ was used at 12 μg ml⁻¹. Immunocomplexes were analysed by SDS-PAGE (12% acrylamide). The protocol for the preparation and analysis of peptides by reversed-phase HPLC has been described²⁰. Briefly, 3–4 × 10⁸ cells were metabolically labelled for 6–8 h with 1 mCi each of [³H]leucine and [³H]lysine. HLA-A2 from detergent extracts was purified using BB7.2 monoclonal antibody³⁶ affinity columns. Eluates were denatured in 10% acetic acid and the dissociated peptides analysed on a reverse-phase column using a Waters HPLC system.

not shown). Thus, the surface expression of apparently stable class I molecules^{29,30} was completely restored in mutant .174 cells, despite the absence of LMP2 and LMP7 (Fig. 3a). In general, the stable assembly of class I molecules is induced by binding of cytosolic peptides of 9 amino acids with specific consensus motifs^{2–4,20,30–32}; longer peptides increase the rate of dissociation of class I heavy chains and β₂m *in vitro* by 2–3 orders of magnitude³². A regular supply of naturally processed peptides was evident from quantitation of the HLA-A2-β₂m complexes after overnight dissociation in cell lysates³⁰: the HLA-A2 heavy chain and β₂m bands were of comparable intensity from purified transfectant C1–C3 cells and .45 cells after precipitation with the MA2.1 monoclonal antibody, which is specific for a conformational epitope on the HLA-A2 heavy chain²⁷. By contrast, most of the HLA-A2 molecules were unstable in .174 cells¹⁸ and their TAP1⁺ or TAP2⁺ transfectant derivatives (Fig. 4a). The effect of TAP1–TAP2 on the stability of the HLA-A2 complexes is thus equivalent to the presence of specific peptides in lysates of .174 cells¹⁸.

To confirm that co-expression of TAP1 and TAP2 was sufficient to restore a normal supply of peptides in .174 cells, endogenously processed peptides associated with HLA-A2 were isolated and analysed by reversed-phase high-performance liquid chromatography (HPLC)²⁰. The peptide profiles from .174 and B6 cells were both very similar to data previously obtained from the .174-derived T2 cells²⁰, including peaks that correspond to peptides from cleaved signal sequences²⁰. The peptides from C3 cells, however, gave a complex profile which was nearly identical to that from wild-type .45 cells (Fig. 4b). Taken together, these results establish that only TAP1 and TAP2 inside the large homozygous deletion in mutant .174 cells are essential for the functional processing of peptides for presentation by class I molecules; LMP2 and LMP7 are not generally required in this pathway. This is in contrast to recent models, in which LMP2 and LMP7 have been directly related to the processing of peptides or to the mechanism of TAP1–TAP2 complexes¹⁵. We emphasize that LMP2 and LMP7 may have a less obvious effect in the B lymphoblastoid cells examined here and possibly a critical function in other cell types. So far, however, there is no evidence to indicate that LMP2 and LMP7 and proteasomes have a role in antigen processing. □

Received 26 August; accepted 13 October 1992.

1. Townsend, A. R. M. *et al.* *Cell* **44**, 959–968 (1986).
2. Townsend, A. *et al.* *Nature* **340**, 443–448 (1989).
3. Falk, K., Rottschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290–296 (1991).
4. Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. *Nature* **353**, 326–329 (1991).
5. Powis, S. J. *et al.* *Nature* **354**, 528–531 (1991).
6. Spies, T. & DeMars, R. *Nature* **351**, 323–324 (1991).
7. Attaya, M. *et al.* *Nature* **355**, 647–649 (1992).
8. Kelly, A. *et al.* *Nature* **355**, 641–644 (1992).
9. Spies, T. *et al.* *Nature* **355**, 644–646 (1992).
10. Monaco, J. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001–3005 (1982).
11. Brown, M. G., Driscoll, J. & Monaco, J. J. *Nature* **353**, 355–357 (1991).
12. Glynne, R. *et al.* *Nature* **353**, 357–360 (1991).
13. Martinez, C. K. & Monaco, J. J. *Nature* **353**, 664–667 (1991).
14. Goldberg, A. L. *Eur. J. Biochem.* **203**, 9–23 (1992).
15. Monaco, J. J. *Immun. Today* **13**, 173–179 (1992).
16. DeMars, R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 8183–8187 (1985).
17. Spies, T. *et al.* *Nature* **348**, 744–747 (1990).
18. Cerundolo, V. *et al.* *Nature* **345**, 449–452 (1990).
19. Kvist, S. & Hamann, U. *Nature* **348**, 446–448 (1990).
20. Wei, M. L. & Cresswell, P. *Nature* **356**, 443–446 (1992).
21. Yang, Y., Waters, J. B., Fruh, K. & Peterson, P. A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4928–4932 (1992).
22. Fruh, K. *et al.* *J. Biol. Chem.* **267**, 22131–22140 (1992).
23. Long, E. O. *et al.* *Hum. Immun.* **31**, 229–235 (1991).
24. Kavathas, P., Bach, F. H. & DeMars, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4251–4259 (1980).
25. Bahram, S., Arnold, D., Bresnahan, M., Strominger, J. L. & Spies, T. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10094–10098 (1991).
26. Haynes, B. F., Reisner, E. G., Hemler, M. E., Strominger, J. L. & Eisenbarth, G. S. *Hum. Immun.* **4**, 273–285 (1982).
27. Ways, J. P., Rothbard, J. B. & Parham, P. *J. Immun.* **137**, 217–222 (1986).
28. Brodsky, F. M., Bodmer, W. F. & Parham, P. *Eur. J. Immun.* **9**, 536–545 (1979).
29. Junggren, H.-G. *et al.* *Nature* **346**, 478–480 (1990).
30. Townsend, A. *et al.* *Cell* **62**, 285–295 (1990).
31. Elliott, T., Cerundolo, V., Elvin, J. & Townsend, A. *Nature* **351**, 402–406 (1991).
32. Cerundolo, V. *et al.* *Eur. J. Immun.* **21**, 2069–2075 (1991).
33. Driscoll, J. & Goldberg, A. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 787–791 (1989).
34. Coligan, J. E., Krusbeek, A. M., Margulies, D. H., Shevach, E. & Strober, W. *Current Protocols in Immunology* (Wiley Interscience, New York, 1991).

35. Ozato, K. & Sachs, D. H. *J. Immunol.* **126**, 317–321 (1980).

36. Brodsky, F. M., Parham, P., Barnstable, C. J., Crumpton, M. J. & Bodmer, W. F. *Immunol. Rev.* **47**, 3–61 (1979).

ACKNOWLEDGEMENTS. We thank R. DeMars for cell lines and M. Surman for technical assistance; E. Long, P. Robbins, J. Strominger and A. Townsend for plasmids and antibodies; and K. Fruh, J. Monaco, A. Townsend and Y. Yang for comments or discussion. D.A. thanks G. Hauptmann for help and encouragement and is supported by the Ministère de la Recherche et de la Technologie. This work was supported by the Cancer Research Institute/Elaine R. Shepard Investigator Award, the Howard Hughes Medical Institute and the NIH.

Proteasome subunits encoded by the major histocompatibility complex are not essential for antigen presentation

Frank Momburg*, Vianney Ortiz-Navarrete*, Jacques Neefjes*, Els Goulmy†, Yvonne van de Wal†, Hergen Spits‡, Simon J. Powis§, Geoffrey W. Butcher§, Jonathan C. Howard§, Peter Walden|| & Günter J. Hämmerling*

*Tumor Immunology Program, German Cancer Research Center, Im Neuenheimer Feld 280, 6900 Heidelberg, Germany

†Department of Immunohaematology and Blood Bank, University Hospital, Bldg 1 E3-Q, PO Box 9600, 2300 RC Leiden, The Netherlands

‡DNAX Research Institute, 901 California Avenue, Palo Alto, California 94304, USA

§Department of Immunology, Agricultural and Food Research Council, Institute of Animal Physiology and Genetics Research, Babraham, Cambridge CB2 4AT, UK

||Max-Planck Institute for Biology, Corrensstrasse 42, 7400 Tübingen, Germany

MAJOR histocompatibility complex (MHC) class I molecules bind and deliver peptides derived from endogenously synthesized proteins to the cell surface for survey by cytotoxic T lymphocytes. It is believed that endogenous antigens are generally degraded in the cytosol, the resulting peptides being translocated into the endoplasmic reticulum where they bind to MHC class I molecules. Transporters containing an ATP-binding cassette encoded by the MHC

TABLE 1 Presentation of the minor histocompatibility antigen HA-2 by T2 cells transfected with transporter genes

Target cells	Effector cells		Anti-HA-2		Q66.9	
	Anti-HLA-A2.1 allo	Anti-HA-2				
	10:1	1:1	10:1	1:1	10:1	1:1
X	71	57	68	43	1	1
Y	72	60	55	26	0	0
T1	96	82	71	64	10	5
T2	100	87	0	1	2	1
T2/TAP1	100	83	0	3	0	1
T2/TAP2	81	61	5	5	3	1
T2/TAP1 + 2	82	75	57	26	0	0

T2 cells transfected with rat *TAP1*^a and *TAP2*^a cDNAs were tested for their capacity to present the endogenously synthesized minor histocompatibility antigen HA-2 to an HLA-A2.1-restricted minor histocompatibility antigen HA-2-specific CTL clone. T2 cells transfected with both transporter genes but not the untransfected T2 could present the HA-2 antigen. An alloreactive CTL clone specific for HLA-A2.1 and the influenza matrix-specific CTL clone Q66.9 (see Table 2a) were included as controls. The HA-2-specific CTL clone, designated HA-2 (ref. 25), and the HLA-A2.1 alloreactive CTL clone²⁶ were mixed with chromium-labelled targets at effector-to-target ratios of 10:1 and 1:1, and specific lysis was measured in a 4-h chromium release assay; values represent per cent specific lysis. X and Y are lymphoblastoid cell lines transformed by Epstein-Barr virus and obtained from HLA-A2.1-positive healthy individuals.

TABLE 2 Presentation of influenza matrix protein M1 by T2 cells transfected with transporter genes

(a)		Q66.9 Effector cells							
		Influenza virus infected				M1 peptide 58–66 added			
Target cells		(expt 1)		(expt 2)		0		1 µg ml ⁻¹	
		5:1	0.5:1	5:1	0.5:1	5:1	0.5:1	5:1	0.5:1
X		28	14	n.t.	n.t.	3	0	23	14
Y		33	15	n.t.	n.t.	0	0	30	12
JY		n.t.	n.t.	31	7	3	0	23	14
T1		50	18	62	17	0	0	68	19
T2		7	4	8	3	7	3	71	21
T2/TAP1		8	3	12	5	6	4	62	31
T2/TAP2		4	1	5	2	6	0	70	25
T2/TAP1 + 2		26	7	33	11	0	0	73	29

(b)		4-30 Effector cells							
		M1-vac infected				M1 peptide 58–66 added			
Target cells		(expt 1)		0		2.5 µg ml ⁻¹		2.5 µg ml ⁻¹	
		10:1	0.6:1	10:1	0.6:1	10:1	0.6:1	10:1	0.6:1
JY		72	32	8	1	80	24		
T1		78	22	11	3	68	24		
T2		4	1	8	0	80	21		
T2/TAP1		14	6	2	2	67	14		
T2/TAP2		7	2	1	0	77	21		
T2/TAP1 + 2		49	16	7	2	75	17		

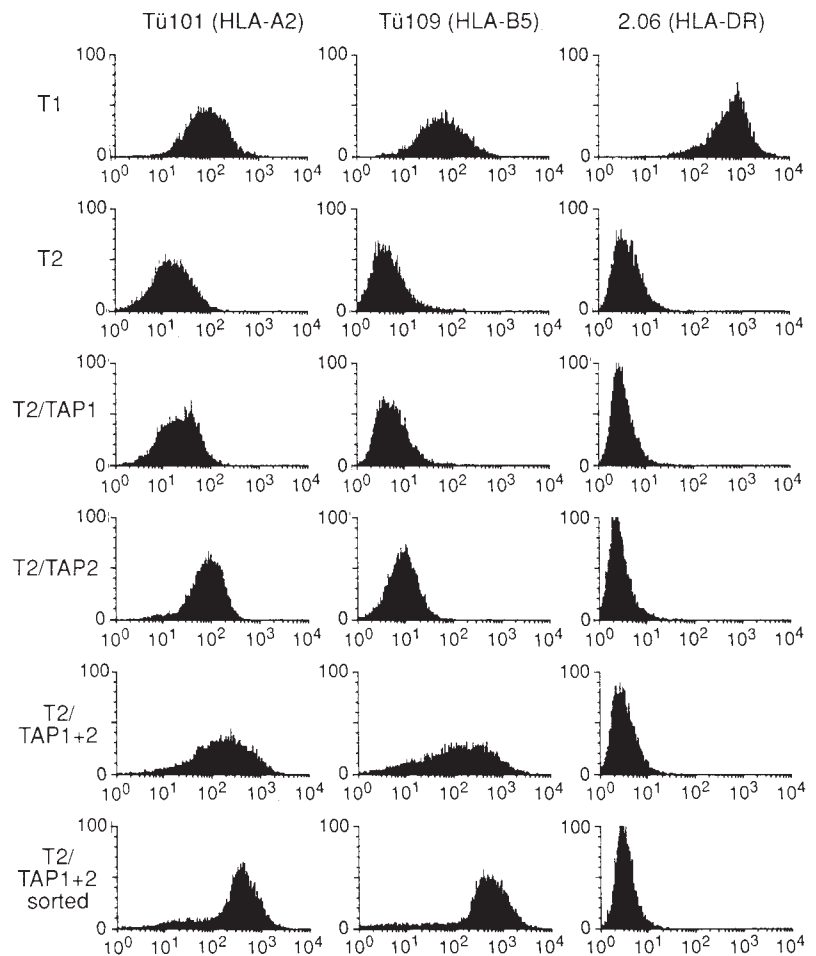
Control cells and T2 cells transfected with rat *TAP1*^a and *TAP2*^a were infected overnight with influenza virus Hongkong 68 and then tested for HLA-A2.1-mediated presentation of the influenza matrix antigen using the CTL clone Q66.9. The results from two representative experiments are shown (a). In additional experiments using recombinant vaccinia virus producing matrix protein (M1-vac) and the M1-specific CTL line 4-30, similar results were obtained (b). Table 2 also shows that all cell lines function as targets after preincubation with exogenously added influenza matrix peptide M58-66 (sequence ILGFVFTLV). (a) Where indicated, target cells were infected overnight with influenza virus Hongkong 68. Q66.9 is a CTL clone recognizing the influenza matrix-derived peptide M58-66 restricted by HLA-A2.1. This clone was raised against the synthetic peptide M58-66 (H.S., unpublished). Lysis was determined in a 4-h chromium-release assay at effector-to-target ratios of 5:1 and 0.5:1. For presentation of exogenously added peptide, target cells were preincubated with M58-66 peptide at 1 µg ml⁻¹ for 2 h before addition of the CTL clone Q66.9. (b) Target cells were infected with M1-vac²⁷ (20 PFU per cell) for 3 h. 4-30 is a fresh CTL line from an HLA-A2.1-positive donor raised against the synthetic peptide M55-73 (P.W., unpublished). Lysis was determined in a 5-h chromium-release assay at effector-to-target ratios ranging from 10:1 to 0.3:1. M58-66 peptide was added as a control 15 min before addition of CTL at 2.5 µg ml⁻¹.

class II region seem to be responsible for this transport^{1–8}. Genes coding for two subunits of the '20S' proteasome (a multicatalytic proteinase) have been found in the vicinity of the two transporter genes in the MHC class II region, indicating that the proteasome could be the unknown proteolytic entity in the cytosol involved in the generation of MHC class I-binding peptides^{9–13}. By introducing rat genes encoding the MHC-linked transporters into a human cell line lacking both transporter and proteasome subunit genes, we show here that the MHC-encoded proteasome subunits are not essential for stable MHC class I surface expression, or for processing and presentation of antigenic peptides from influenza virus and an intracellular protein.

The 20S proteasome consists of about 20–30 subunits with *M_r*s between 15,000 and 30,000 that are encoded by distinct sets of genes¹⁴. We used the human lymphoblastoid B cell-derived mutant T2 line to investigate whether or not the MHC-linked proteasome subunits are essential for MHC class I expression and antigen presentation. T2 has a large homozygous deletion of the MHC class II region¹⁵, which encompasses the genes for the two ATP-binding cassette (ABC) transporter polypeptides, TAP1 and TAP2, and the two genes for the MHC-encoded proteasome subunits, Lmp2 and Lmp7. In the absence of transporters, most HLA class I molecules are devoid of the peptides

FIG. 1 Restoration of HLA class I expression on T2 cells transfected with transporter genes. The cDNAs for the two rat transporter chains *TAP1a* and *TAP2a* were used individually or as a mixture to transfect T2 cells. HLA cell-surface expression was measured by cytofluorometry. In five independent cotransfections with *TAP1* and *TAP2* we always obtained a substantial fraction of cells expressing high levels of HLA class I, similar to the example shown here.

METHODS. In each experiment, 10^7 T2 cells were transfected by electroporation with $\sim 2 \mu\text{g}$ each of rat cDNAs *TAP1^a* (*mtp1^a*, clone 510-15) or *TAP2^a* (*mtp2^a*, clone 441-11) in the pH β APr-1-neo expression vector^{1,5,21}. Both plasmids contain the β -actin promoter and the neomycin-resistance gene. Selection was in 1 mg ml^{-1} G418 (Gibco). After 4 to 6 weeks, cytofluorometry was done with the bulk cultures using a FACScan (Becton and Dickinson). Antibodies used were Tü101 against HLA-A2 (ref. 28), Tü109 against HLA-Bw4 (ref. 29), and 2.06 against HLA-DR, -DP and -DQ (ref. 30), detected by fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG (Sigma). Staining with antibodies Tü48 (HLA-Bw4; ref. 31) and BB7.2 (HLA-A2; ref. 32) yielded comparable results, except that BB7.2 stained HLA-A2 antigens on untransfected T2 cells more strongly than antibody Tü101. The unseparated bulk culture of T2 cotransfected with both transporters yielded a broad peak containing both negative and positive cells. Bottom row shows staining profiles of T2/TAP1+2 enriched by cell sorting for high expression of HLA-B5 using antibody Tü109 and FACStar Plus cell sorter.



necessary for stable assembly of the class I heavy chain with β_2 -microglobulin and for expression at the cell surface¹⁶⁻¹⁸. Therefore the T2 cell line shows strongly decreased expression of HLA-B5 and partially decreased expression of HLA-A2.1 molecules and is deficient in antigen presentation^{19,20}.

We transfected the T2 cells with the rat ABC transporter complementary DNAs rat *TAP1^a* and rat *TAP2^a*, previously named *mtp1^a* and *mtp2^a*, respectively^{1,5,21}. The resulting transfectants were analysed for class I cell-surface expression. Parental T1 cells, from which T2 cells were derived, were positive for HLA-A2.1 and -B5 and also for HLA class II molecules, whereas T2 cells did not stain for HLA-B5 and only weakly for HLA-A2.1 (Fig. 1). Transfection with rat *TAP1^a* alone did not alter HLA expression. Transfection with *TAP2^a* resulted in a slight increase of HLA-A2 and HLA-B5 expression, which we are investigating at present. But when both *TAP1^a* and *TAP2^a* rat transporter genes were transfected into T2 cells, HLA-A2.1 and HLA-B5 expression was restored to levels two to three times higher than on T1 cells. Expression of class I in this bulk culture was improved by sorting for cells expressing high levels of HLA-B5 (bottom panel).

Western blots using sera raised against rat TAP1 and TAP2 showed that the T2/TAP1+2 transfectant expressed the TAP polypeptides in amounts comparable to rat strain PVG.R19 lymphoblasts stimulated by concanavalin A (Fig. 2a). The absence of the MHC-encoded proteasome subunits in the T2 transfectants was confirmed by immunoprecipitation (not shown) and messenger RNA analysis (Fig. 2b). It is possible that in the absence of Lmp2 and Lmp7, peptides of inappropriate length are generated, which would result in a decreased stability of the assembled class I molecules. Lysates of biosynthetically labelled cells were incubated at 37 °C for different times. Class

I molecules were then immunoprecipitated with monoclonal antibody W6/32, which only recognizes class I molecules associated with β_2 -microglobulin. Class I molecules devoid of peptides are unstable under these conditions and lose the epitope recognized by W6/32 (ref. 22). HLA-A2 and B5 were separated by one-dimensional isoelectric focusing (Fig. 2c). Most HLA-A2 molecules expressed in T2 cells are unstable to exposure at 37 °C, whereas most HLA-A2 molecules in T1 cells and T2 cells reconstituted with TAP1 and TAP2 are stable over a 4-h exposure to 37 °C. A more drastic effect of temperature on the stability of HLA-B5 is evident. HLA-B5 in T2 cells is unstable at 37 °C, whereas it is stable both in T1 cells and in T2/TAP1+2 cells. Thus, expression of TAP1 and TAP2 in T2 cells results in proper stabilization of both HLA-A2 and -B5 molecules. These results demonstrate that expression of transporter polypeptides alone, in the absence of the MHC-encoded proteasome subunits, is sufficient for apparently normal and stable class I expression.

Next, T2 cells transfected with rat *TAP1^a* and rat *TAP2^a* were investigated for their ability to process and present endogenous proteins. For this purpose T2 transfectants were tested with an HLA-A2.1-restricted cytotoxic T lymphocyte (CTL) clone recognizing a minor histocompatibility antigen (HA-2). Control cells, including T1, were lysed, but not T2 cells or T2 transfectant with only one transporter gene (Table 1). In contrast, T2 cells transfected with both transporter genes were lysed efficiently. An HLA-A2.1 alloreactive CTL clone included as a positive control strongly lysed all transfectants, including the parental T2 cells. As the HLA-A2.1 molecules on T2 cells appear to carry only a limited variety of peptides derived from signal sequences^{23,24}, our data suggest that the anti-HLA-A2.1 CTL clone recognizes such a signal sequence-derived peptide.

Because the nature of the HA-2 antigen is not yet known, we

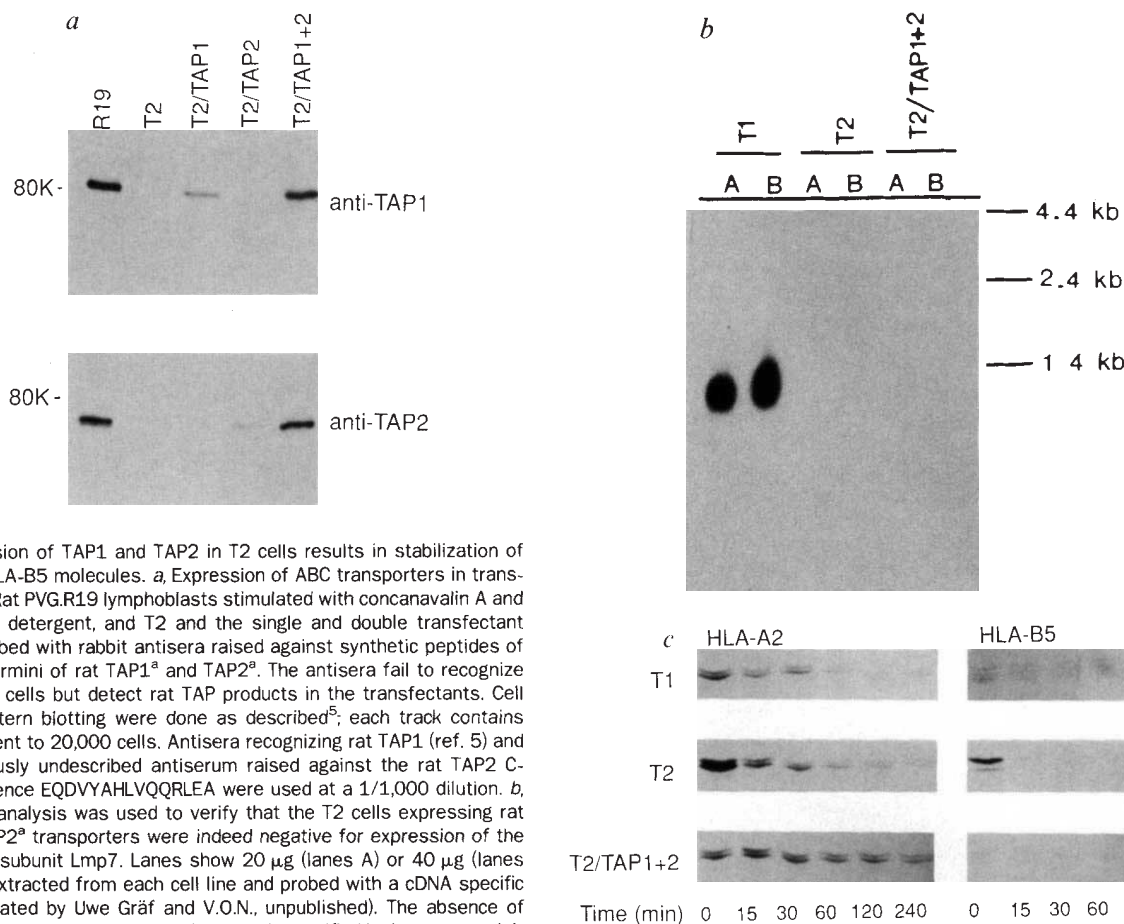


FIG. 2 Expression of TAP1 and TAP2 in T2 cells results in stabilization of HLA-A2 and HLA-B5 molecules. *a*, Expression of ABC transporters in transfectant cells. Rat PVG.R19 lymphoblasts stimulated with concanavalin A and lysed in NP40 detergent, and T2 and the single and double transfectant cells were probed with rabbit antisera raised against synthetic peptides of the carboxy termini of rat TAP1^a and TAP2^b. The antisera fail to recognize material in T2 cells but detect rat TAP products in the transfectants. Cell lysis and western blotting were done as described⁵; each track contains lysate equivalent to 20,000 cells. Antisera recognizing rat TAP1 (ref. 5) and a new, previously undescribed antiserum raised against the rat TAP2 C-terminal sequence EQDVYAHVQQRLEA were used at a 1/1,000 dilution. *b*, Northern blot analysis was used to verify that the T2 cells expressing rat TAP1^a and TAP2^b transporters were indeed negative for expression of the MHC encoded subunit Lmp7. Lanes show 20 μ g (lanes A) or 40 μ g (lanes B) total RNA extracted from each cell line and probed with a cDNA specific for Lmp7 (isolated by Uwe Gräf and V.O.N., unpublished). The absence of the MHC-encoded proteasome subunits was also verified by immunoprecipitation (not shown) using a rabbit serum against the 20S proteasome¹¹. *c*, Stability of MHC molecules was assayed by exposing cell lysates to 37 °C for different times. Stable class I molecules were recovered with the monoclonal antibody W6/32. Both HLA-A2 and -B5 molecules are stabilized by expressing TAP1 and TAP2 in T2 cells and regain a similar stability as class I molecules in T1 cells. We labelled 14×10^6 T1, T2 or T2/TAP1 + 2 cells for 15 min with 200 μ Ci ³⁵S-methionine and cysteine, respectively.

analysed the processing of a well-defined antigen, influenza virus matrix protein. Cells were infected overnight with influenza virus and then used as targets for the CTL clone Q66.9, which recognizes the influenza matrix-derived peptide epitope M58-66 restricted by HLA-A2.1. Representative experiments in Table 2a show that infected T1 cells, the HLA-A2.1-positive cell line JY, and X and Y targets were killed, but not T2 cells or T2 transfected with the individual transporter genes. In contrast, the T2/TAP1 + 2 cells expressing both transporter polypeptides were lysed. This lysis (26 and 33% at a 5:1 effector to target ratio) was weaker than the lysis of T1 cells (50% and 62%, respectively) but with the other influenza virus-infected cell lines, JY, X and Y, only 28 to 33% lysis was again obtained. Results were comparable when target cells were infected with a recombinant vaccinia virus making the influenza M1 matrix protein (M1-vac) and the M1-specific CTL line 4-30 was used (Table 2b). Lysis of M1-vac-infected T2/TAP1 + 2 cells was again less than lysis of M1-vac-infected T1 cells. All transfectants were able to present the synthetic matrix peptide M58-66 when it was added exogenously. It is clear from these data that the capacity to form and present this influenza epitope has been returned to the T2/TAP1 + 2 transfectant, although it is possible that this capacity is suboptimal.

We conclude that the heterodimeric MHC-encoded transporter is alone sufficient for MHC class I-mediated antigen presentation, at least for the two antigens tested here. Thus, there

Cells were lysed in NP40 lysis mixture and their nuclei removed. Lysates were precleared with normal rabbit serum and equal amounts incubated at 37 °C for the times indicated. After preclearing again with normal rabbit serum, class I molecules were immunoprecipitated with monoclonal antibody W6/32. Immunoprecipitates were analysed by one-dimensional isoelectric focusing²²; only HLA-A2 and -B5 molecules are shown.

does not seem to be an absolute requirement for the MHC-linked proteasome subunits Lmp2 and Lmp7. Our observations could mean that the presence of genes encoding two of the proteasomes in the MHC may be fortuitous and that the 20S proteasome does not play a role in antigen presentation. It is also possible however that the proteasome is only one of several proteolytic enzymes with the capacity to digest cytosolic antigen into class I-binding peptides. A proteasome lacking the MHC-encoded subunits may well have adequate proteolytic activity and could be the source of peptides assembled in our T2/TAP1 + 2 transfectant. The MHC-encoded proteasome subunits, which are inducible by interferon- γ , may nevertheless function to increase the overall proteolytic activity or the range of peptides generated by the proteasome from antigens in the context of an immune response, to a virus for example. Investigation of T2 cells restored with transporters and of the MHC-encoded proteasome subunits will help to clarify this issue. □

Received 5 June; accepted 17 September 1992.

- Deverson, E. V. *et al.* *Nature* **348**, 738–741 (1990).
- Trowsdale, J. *et al.* *Nature* **348**, 741–744 (1990).
- Spies, T. *et al.* *Nature* **348**, 744–747 (1990).
- Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723–1726 (1990).
- Powis, S. J. *et al.* *Nature* **354**, 528–531 (1991).
- Kelly, A. *et al.* *Nature* **355**, 641–644 (1992).
- Spies, T. *et al.* *Nature* **355**, 644–646 (1992).
- Attaya, M. *et al.* *Nature* **355**, 647–649 (1992).
- Glynne, R. *et al.* *Nature* **353**, 357–360 (1991).

10. Brown, M. G., Driscoll, J. & Monaco, J. J. *Nature* **353**, 355-357 (1991).
11. Ortiz-Navarrete, V. *et al. Nature* **353**, 662-664 (1991).
12. Martinez, C. K. & Monaco, J. J. *Nature* **353**, 664-667 (1991).
13. Kelly, A. *et al. Nature* **353**, 667-668 (1991).
14. Rivett, A. J. *Archs Biochem. Biophys.* **268**, 1-8 (1989).
15. Salter, R. D., Howell, D. N. & Cresswell, P. *Immunogenetics* **21**, 235-235 (1985).
16. Townsend, A. *et al. Cell* **62**, 285-295 (1990).
17. Kvist, S. & Hamann, U. *Nature* **348**, 446-448 (1990).
18. Ortiz-Navarrete, V. & Hämmerling, G. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3594-3597 (1991).
19. Cerundolo, V. *et al. Nature* **345**, 449-452 (1990).
20. Hosken, N. A. & Bevan, M. J. *Science* **248**, 367-370 (1990).
21. Powis, S. J. *et al. Nature* **357**, 211-215 (1992).
22. Baas, E. J. *et al. J. exp. Med.* **176**, 147-156 (1992).
23. Henderson, R. A. *et al. Science* **255**, 1264-1266 (1992).
24. Wei, M. L. & Cresswell, P. *Nature* **356**, 443-446 (1992).
25. Goulmy, E. *Transplant. Rev.* **2**, 29-53 (1988).
26. Horai, S., v.d. Poel, J. & Goulmy, E. *Immunogenetics* **16**, 135-142 (1992).
27. Gotch, F., Rothbard, J., Howland, K., Townsend, A. & McMichael, A. *Nature* **326**, 881-882 (1987).
28. Müller, C., Shi, L., Schneider, M., Ziegler, A. & Wernet, P. *Hum. Immun.* **6**, 189-197 (1983).
29. Müller, C. *et al. Hum. Immun.* **14**, 333-349 (1985).
30. Charron, D. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6567-6571 (1979).
31. Müller, C., Ziegler, A., Müller, G., Schunter, F. & Wernet, P. *Hum. Immun.* **5**, 269-281 (1981).
32. Brodsky, F. M., Parham, P., Barnstable, C. J., Crumpton, M. J. & Bodmer, W. F. *Immunol. Rev.* **47**, 3-61 (1979).

ACKNOWLEDGEMENTS. We thank M. Post for technical help, A. Arzberger for cell sorting, A. Ziegler for antibodies, V. Cerundolo for T1 and T2 cells, A. McMichael for M1-vac recombinant vaccinia virus, U. Esslinger for preparation of the manuscript, and C. M. Mellef for discussion. This work was partially supported by the Deutsche Forschungsgemeinschaft. S.J.P. is supported by the UK Arthritis and Rheumatism Council; J.N. is supported by an EMBO fellowship.

Abrogation by *c-myc* of G1 phase arrest induced by *RB* protein but not by p53

David W. Goodrich & Wen-Hwa Lee

Center for Molecular Medicine, Institute of Biotechnology,
University of Texas Health Science Center, 15355 Lambda Drive,
San Antonio, Texas 78245, USA

INACTIVATING mutations of the retinoblastoma gene (*RB*) are found in a wide variety of tumour cells¹. Replacement of wild-type *RB* can suppress the tumorigenicity of some of these cells, suggesting that the *RB* protein (Rb) may negatively regulate cell growth²⁻⁴. As activation of *c-myc* expression promotes cell proliferation and blocks differentiation, it may positively regulate cell growth⁵⁻⁹. The *c-myc* protein is localized in the nucleus and can physically associate with *RB* protein *in vitro*¹⁰, hence *c-myc* may functionally antagonize *RB* function. Microinjection of Rb in G1 phase reversibly arrests cell-cycle progression¹¹. Here we co-inject *RB* protein with *c-myc*, *EJ-ras*, *c-fos* or *c-jun* protein. Co-injection of *c-myc*, but not *EJ-ras*, *c-fos* or *c-jun*, inhibits the ability of Rb to arrest the cell cycle. The *c-myc* does not inhibit the activity of another tumour suppressor, p53 (ref. 12). Thus, *c-myc* and *RB* specifically antagonize one another in the cell.

Various forms of partially purified *c-myc*-glutathione S-transferase (GST) fusion protein were co-injected with an amino-truncated form of *RB* protein (p56RB)¹¹. p56RB contains the domains required for binding several cellular proteins, including *c-myc*¹³⁻¹⁵. After injection in early G1 phase of the cell cycle, cells were incubated in the presence of bromodeoxyuridine (Br-dU); cells entering S phase will incorporate Br-dU and can be stained with an anti-Br-dU antibody. The percentage of injected cells stained with Br-dU was used to determine the effect of injected proteins on G1 phase progression¹¹.

Full-length *c-myc* fusion protein abolishes cell-cycle arrest induced by injection of p56RB (Table 1). Deletion of the amino terminus (GSTmyc98-439 and GST49-439), the carboxy terminus (GSTmyc1-262 and GSTmyc1-391) or internal amino acids (GSTmycΔ235-262) from the full-length fusion reduces the ability of *c-myc* to counteract p56RB (Table 1). An insertion mutation generating a protein with four additional amino acids between 234 and 235 of *c-myc* (GSTmycI234) also has a reduced ability to inhibit p56RB activity. The *myc* fusion proteins do not prevent Br-dU incorporation if injected alone (Table 1).

To determine whether the functional interaction between p56RB and *c-myc* protein is specific, we co-injected p56RB with purified *EJ-ras*, *c-fos* or *c-jun* proteins¹⁶⁻¹⁸. G1 phase arrest induced by p56RB was not relieved by co-injection with either *EJ-ras*, *c-fos* or *c-jun* protein (Table 1). Although *EJ-ras* does not inhibit p56RB activity, the cells show a distinct morphology characteristic of injection with *EJ-ras* protein alone (data not shown). Further, the *c-fos* and *c-jun* protein preparations, when mixed together, retain the ability to bind an oligonucleotide containing the AP-1 site in a DNA-mobility shift assay (data not shown)¹⁸. Therefore the protein preparations were functional but did not inhibit p56RB activity. Injection of *EJ-ras*, *c-fos* or *c-jun* alone did not prevent incorporation of Br-dU.

In contrast to an SV40 T-antigen peptide¹¹, *c-myc* binding alone was probably not sufficient to abrogate Rb-induced cell cycle arrest. The ratio of injected *c-myc* fusion protein to p56RB was much less than one (Fig. 1a), so only a fraction of p56RB would be in complex with *myc* protein. Further, all of the GST-*myc* fusion proteins tested, with the possible exception of GSTmycΔ235-262, could bind purified full-length Rb protein *in vitro* (Fig. 1b). The mutant GSTmyc protein preparations, except GSTmyc1-390, could also complex with *max* protein (Max) (Fig. 1c)¹⁹. Although most of the mutant GST-*myc* proteins could bind Rb and Max, suggesting that both the amino-terminal Rb-binding domain and the carboxy terminal Max

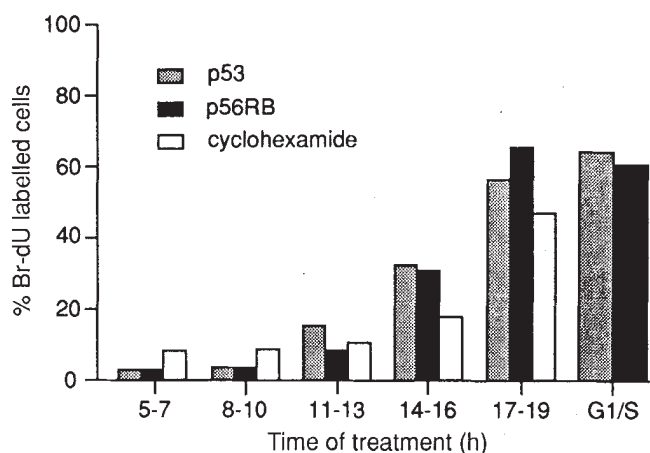


FIG. 1. a, Aliquots of p56RB (5 μ l, lane 1), p53 (25 μ l, lane 2), GSTmyc (25 μ l, lane 3), and SV40 T antigen (20 μ l, lane 4) used for microinjection were resolved by SDS-PAGE and stained with Coomassie blue. b, GST-SV40 T antigen (lane 1), GSTmycΔ235-262 (lane 2), GSTmycI234 (lane 3), GSTmyc1-390 (lane 4), GSTmyc98-439 (lane 5), GSTmyc49-439 (lane 6), GSTmyc (lane 7) and GST (lane 8) fusion proteins bound to affinity resin were mixed with p110RB, the bound proteins eluted and resolved by SDS-PAGE, western blotted, and the blot developed with antibodies directed against *RB* protein. c, GSTmyc (lane 1), GSTmyc49-439 (lane 2), GSTmycΔ235-262 (lane 3), GSTmyc98-439 (lane 4), GSTmyc1-390 (lane 5), GSTmycI234 (lane 6), or GST (lane 7) fusion proteins bound to affinity resin were mixed with purified *max* protein and processed as in b using a polyclonal antibody directed against Max.

METHODS. SV40 T antigen was expressed in a baculovirus system and purified on an anti-T antigen pAb419 immunoaffinity column³². Approximately equal amounts of GSTmyc fusion proteins bound to glutathione beads were mixed with 50 ng of p110RB in a buffer containing 20 mM Tris, pH 8, 200 mM NaCl, 1 mM EDTA, 0.1 μ g ml⁻¹ BSA, and 0.5% NP40. Beads were incubated for 20 min at 4°C, washed with binding buffer, the proteins eluted with SDS-PAGE loading buffer, resolved by 10% SDS-PAGE, electroblotted, and developed with anti-Rb monoclonal antibody 11D7 and alkaline phosphatase-conjugated anti-mouse IgG. The p110RB protein was expressed and purified from *E. coli* like p56RB (C. Hensley, unpublished). Alternatively, GSTmyc fusion protein affinity beads were mixed with 50 ng Max in buffer containing 20 mM HEPES, pH 7.2, 50 mM NaCl, 0.1% NP40, 0.1 μ g ml⁻¹ BSA and 5 mM dithiothreitol. Incubation was continued for at least 2 h at 4°C before proteins were eluted with loading buffer, resolved by 15% SDS-PAGE, blotted, and developed using an anti-Max polyclonal antibody¹⁹.

TABLE 1 Co-injection of synchronous Saos-2 cells in early G1

Proteins	Cells injected	Br-dU-labelled cells	% Entering S phase
p56RB + buffer	635	19	3
+ GSTmyc	453	223	49
+ GSTmyc1-262	599	6	1
+ GSTmyc1-390	671	41	6
+ FSTmyc49-439	490	30	6
+ GSTmyc98-439	339	6	2
+ GSTmycΔ235-262	522	9	2
+ GSTmyc1234	242	14	6
+ EJ-ras	501	35	6
+ c-fos	148	5	3
+ c-jun	433	7	2
+ fos/jun	157	8	5
GSTmyc	245	122	50
EJ-ras	525	224	43
c-fos	351	157	45
c-jun	371	192	52
GSTmyc1-262	271	121	45
GSTmyc98-439	302	168	56
Histone	553	312	56

GSTmyc represents the full-length *c-myc* protein fused to a domain of glutathione S-transferase. The names of the *myc* mutants indicate the amino-acid numbers of the full-length coding sequence that are included within the protein. GSTΔmyc235-262 describes the numbers of the amino acids deleted, in frame, from the full-length *myc* fusion protein. GSTmyc1234 indicates a protein containing the amino acids DELA inserted between amino-acid number 234 and 235. The results are compiled from at least three independent experiments. Proteins were microinjected into synchronized Saos-2 osteosarcoma cells which lack both wild-type *RB* and p53 (refs 21, 22, 31). Cells were synchronized, injected, labelled and stained as described¹¹. The concentration of p56RB preparations was 0.5–1.5 mg ml⁻¹ as determined by absorbance at 280 nm and Coomassie blue staining after SDS-PAGE. EJ-ras, c-fos and c-jun proteins were used at concentrations of 0.05 to 0.3 mg ml⁻¹; their purification has been described elsewhere^{16–18}. The concentration of the GSTmyc preparation has been compared to p56RB by SDS-PAGE (Fig. 1). The mutated *myc* fusion proteins were used at a concentration equal to or greater than GSTmyc as determined by SDS-PAGE (data not shown). Fusion proteins were expressed and purified from *Escherichia coli* as described¹⁰. Histone was purchased commercially (Boehringer Mannheim) and dialysed against injection buffer. All protein preparations contained a biotinylated rabbit anti-goat antibody at 1 mg ml⁻¹ which was used to mark injected cells. Injection of this antibody alone had no detectable effect on cells. All injections were done between 5 and 9 h after release from nocadazole arrest.

binding domain were structurally intact, they were all defective in antagonizing Rb function (Table 1). We cannot rule out the possibility, however, that some *myc* mutants lack activity in the microinjection assay because of technical limitations in their purification.

We have also co-injected p53 with *c-myc* protein. GSTmyc does not inhibit the activity of p53 (Table 2). Like p56RB, injection of wild-type p53 in early G1 prevents entry into the following S phase (Table 2). The p53 block to S phase entry is specific, because co-injection with an anti-p53 antibody abolishes activity. Like p56RB¹¹, inhibition of S phase entry by p53 is reversible because an increasing percentage of cells enter S phase if incubation with Br-dU is continued for 2–3 days (Table 2). Both p53 and p56RB preparations bind purified SV40T antigen *in vitro* (data not shown), and T antigen can suppress the activity of injected p53. The time in G1 phase during which Saos-2 cells are susceptible to arrest by p53 is also similar to the time defined for p56RB (ref. 11). A transition in the response of cells to p53 injection is observed 8–10 hours before the normal onset of S phase (Fig. 2). p53 is also inactive if injected at the aphidicolin arrest point. Therefore the block to entry of S phase is not due entirely to direct inhibition of DNA synthesis. In contrast to p56RB (ref. 11), p53 does not arrest G1 progression upon injection into two cell lines (CV-1 and U2OS) that already expressed wild-type p53 (Table 2) (P. L. Chen, personal communication)^{20–22}. Although the activity of p56RB and p53 appear to be similar, they interact differently with *c-myc* and cell line CV-1 or U2OS.

Our studies have indicated that the *c-myc* and *RB* gene products functionally antagonize one another in the cell. This functional interaction is specific as the activity of another tumour-suppressor gene product, p53, is unaffected by *myc*, and because other oncogene proteins do not affect the activity of p56RB. Our results are consistent with a preliminary report that co-transfection of *c-myc* and *RB* prevents inhibition of cell growth normally associated with transfection of *RB* alone (K. G. Wiman and G. Klein, unpublished). We suggest that *c-myc* may be one of a number of genes that function downstream of *RB* because *c-myc* is dominant to *RB* in our microinjection assay. Consistent with this notion, transforming growth factor-β downregulation of *c-myc* transcription is blocked by viral-transforming proteins that bind Rb, suggesting that Rb may be an intermediary in regulation of *c-myc* by TGF-β (refs 23, 24).

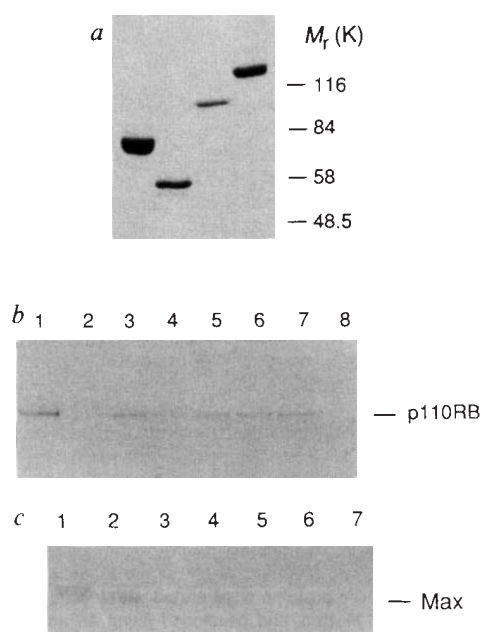


FIG. 2 Injection of p53 or p56RB into Saos-2 cells at different times throughout G1 to S phase. Saos-2 cells are injected with either p53 (speckled bar) or p56RB (black bar) at the indicated times, in hours, after release from mitotic arrest. 'G1/S' represents injection of cells arrested by aphidicolin. The percentage of injected cells progressing into S phase during the incubation period is recorded. Alternatively, cyclohexamide is added to synchronized Saos-2 cells at the indicated times after release from mitotic arrest; the percentage of cells entering S phase during the labelling period is indicated by the unshaded bars. The results are the average of at least three independent experiments.

METHODS. Synchronization of cells, microinjection, labelling with Br-dU, and antibody staining were all as described¹¹. Cyclohexamide was used at a concentration of 5 µg ml⁻¹. Cells were labelled with Br-dU until 32–36 h after release from mitotic arrest.

TABLE 2 Microinjection of synchronous cells in early G1 with p53

Cell line	Protein	Cells injected	Br-dU-labelled cells	% Entering S phase	Labelling time
Saos-2	p53	413	12	3	24
	+buffer	573	54	9	24
	+ α p53 421	584	298	51	24
	+Tag	288	207	72	24
	+GSTmyc	227	2	1	24
	p53	245	45	18	48
	p53	244	149	61	72
CV-1	Histone	553	312	56	24
	p53	332	176	53	24
U2OS	p53	168	144	86	24

Cells were synchronized, injected, labelled and stained as described in Table 1 legend. Results are compiled from at least three independent experiments for Saos-2 and CV-1 cells, and from two experiments for U2OS cells. The p53 protein was made with a baculovirus expression vector and purified from insect cells by immunoaffinity chromatography (J. Duncan, manuscript in preparation). The concentration of p53 has been compared to p56RB and SV40 T antigen in Fig. 1. The anti-p53 antibody was purified by immunoaffinity chromatography from hybridoma culture supernatants³². Histone was purchased commercially (Boehringer Mannheim) and dialysed against injection buffer.

Consistent with the notion that *RB* and *c-myc* participate in the same or overlapping regulatory networks, tissues affected by the absence of *RB* overlap the tissues affected by overexpression of *c-myc*. The phenotype of homozygous *RB* null mice includes embryonic lethality, with defects in haematopoietic and neuronal development^{25,33,34}. In contrast, homozygous p53 null mice develop normally but have a predisposition to tumours when they are older²⁶. Interestingly, overexpression of *myc* also interferes with haematopoietic differentiation *in vitro*^{5-7,9,27} and *in vivo*^{28,29}. As mutations throughout the *c-myc* coding sequence apparently affect its activity in our assay, the regions of *c-myc* required to inhibit p56RB correlate most closely with those required to inhibit differentiation^{29,30}.

Given the effects of *RB* and *c-myc* on cell-cycle progression, it is conceivable that they are involved in regulating the cell's exit from the cell cycle during differentiation. To confirm this hypothesis, it will be important to characterize the regulatory networks involving *RB*. The co-injection assay presented here can be used to identify the members of these networks. Attractive candidates include the proteins that can physically associate with *RB*. □

Received 6 July; accepted 28 September 1992.

- Bookstein, R. & Lee, W.-H. *CRC Crit. Rev. Oncogenesis* **2**, 211-227 (1991).
- Bookstein, R., Shew, J.-Y., Chen, P.-L., Scully, P. & Lee, W.-H. *Science* **247**, 712-715 (1990).
- Goodrich, D. W., Chen, Y., Scully, P. & Lee, W.-H. *Can. Res.* **52**, 1968-1973 (1992).
- Huang, H.-J. S. *et al. Science* **242**, 1563-1566 (1988).
- Bar-Ner, M., Messing, L. T., Cultraro, C. M., Birrer, M. J. & Segal, S. *Cell Growth Differ.* **3**, 183-190 (1992).
- Copolla, J. S. & Cole, M. D. *Nature* **320**, 760-763 (1986).
- Dmitrovsky, E. *et al. Nature* **322**, 748-750 (1986).
- Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-602 (1983).
- Prochownik, E. V. & Kukowska, J. *Nature* **322**, 848-850 (1986).
- Rustgi, A. K., Dyson, N. & Bernards, R. *Nature* **352**, 541-544 (1991).
- Goodrich, D. W., Wang, N. P., Qian, Y.-W., Lee, E. Y.-H. P. & Lee, W.-H. *Cell* **67**, 293-302 (1991).
- Levine, A. J., Momand, J. & Finlay, C. A. *Nature* **351**, 453-456 (1991).
- Huang, S., Lee, W.-H. & Lee, E. Y.-H. P. *Nature* **351**, 1211-1217 (1991).
- Kaelin, W. G. J., Pallas, D. C., DeCaprio, J. A., Kaye, F. J. & Livingston, D. M. *Cell* **64**, 521-532 (1991).
- Lee, W.-H. *et al. Cold Spring Harbor quant. Biol.* **LVI**, 211-217 (1991).
- Kamata, T. & Kung, H.-F. *Molec. cell. Biol.* **10**, 880-886 (1990).
- Stacey, D. W. & King, H.-F. *Nature* **310**, 508-511 (1984).
- Xanthoudakis, S. & Curran, T. *EMBO J.* **11**, 653-665 (1992).
- Blackwood, E. M. & Eisenman, R. N. *Science* **251**, 1211-1217 (1991).
- Diller, L. *et al. Molec. cell. Biol.* **10**, 5772-5781 (1990).
- Masuda, H., Miller, C., Koeffler, H. P., Battifora, H. & Cline, M. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7716-7719 (1987).
- Miller, C. W. *et al. Cancer Res.* **50**, 7950-7954 (1990).
- Pietenpol, J. A. *et al. Cell* **61**, 777-785 (1990).
- Pietenpol, J. A., Munger, K., Howley, P. M., Stein, R. W. & Moses, H. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10227-10231 (1991).

- Lee, E. Y.-H. P. *et al. Nature* **359**, 288-294 (1992).
- Donehower, L. A. *et al. Nature* **215**-221 (1992).
- Kume, T. U., Takada, S. & Obinata, M. *J. molec. Biol.* **202**, 779-786 (1988).
- Bishop, J. M. & Varmus, H. in *RNA Tumor Viruses* (eds Weiss, R., Teich, N., Varmus, H. & Coffin, J.) 999-1108 (Cold Spring Harbor Laboratory Press, New York, 1982).
- Cole, M. D. *A. Rev. Genet.* **20**, 361-384 (1986).
- Predergast, G. C. & Ziff, E. B. *Trends Genet.* **8**, 91-96 (1992).
- Shew, J.-Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 6-10 (1990).
- Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. *J. Virol.* **39**, 861-869 (1981).
- Jacks, T. *et al. Nature* **359**, 295-300 (1992).
- Clarke, A. R. *et al. Nature* **359**, 328-330 (1992).

ACKNOWLEDGEMENTS. We thank P.-L. Chen for making some of the GSTmyc constructions and the baculovirus clone used for production of p53; C. Hensley for purifying p56RB and p110RB; J. Duncan for purified p53; S. Baker and T. Curran for *c-fos* and *c-jun* proteins; H.-F. Kung for EJ-ras protein; and D. Ayer and R. Eisenman for *max* protein and *max*-specific antibody. D.W.G. is a fellow of the Leukemia Society of America. This research is supported by grants to W.-H. Lee from NIH and the Council for Tobacco Research.

Actin cables and epidermal movement in embryonic wound healing

Paul Martin* & Julian Lewis†

* Department of Human Anatomy, Oxford University, Oxford OX1 3QX, UK
† ICRF Developmental Biology Unit, Department of Zoology, Oxford University, Oxford OX1 3PS, UK

SKIN wounds in embryos heal rapidly and perfectly. Even though the epidermis appears to be stretched taut over the surface of a structure such as a growing limb bud, its response to wounding is to close over the lesion, rather than to gape more widely. In adult wounds, the epidermis seems to migrate by means of lamellipodia, crawling over the exposed connective tissue¹⁻⁵. But in embryonic wounds we do not see lamellipodia. The epidermis at the edge of the wound looks smooth, as though under a circumferential tension. Here we show that a cable of filamentous actin appears to run continuously around most of the wound margin. It is confined to the single row of basal cells at the free edge of the epidermis. We suggest that the actin cable acts as a contractile 'purse string' to close up the embryonic wound.

We studied the healing of a standardized lesion on the dorsal surface of a four-day chick embryonic wing bud. The epidermis at this stage consists simply of a single basal layer of cuboidal cells overlaid by an outer layer of squamous cells, called the periderm. The lesion is made by dissecting away, using a tungsten needle, a square patch of embryonic skin measuring roughly 0.5 × 0.5 mm and about 0.1 mm deep (Fig. 1a). The time course of healing can be observed in the egg, from camera lucida drawings of living specimens followed throughout the process. We found that the wound margin moves in at a steady rate of 10-15 μ m per hour, and the wound is closed within 18 to 24 hours. In adult wounds, closure is brought about in part by contraction of the wound connective tissue (the granulation tissue)⁶. It was therefore important to check that in embryonic wounds the epidermis was actively moving over the mesenchyme and not merely riding passively on mesenchymal contraction. To test this, we marked the boundaries of the exposed mesenchyme immediately after wounding by applying spots of the lipophilic dye DiI, which, once incorporated, remains confined to the cells that have taken it up⁷. Wounds were then left to heal for 18 hours and the position of the epithelial wound margin was noted in relation to the mesenchymal markers (Fig. 2). The mesenchyme does seem to contract to a modest extent, but more importantly, the epidermis clearly moves inwards over the marked mesenchyme cells. We conclude that this movement must be driven by forces generated in the epidermis.

Scanning electron microscopy gives some insight into the behaviour of the epithelium at the wound margin (Fig. 1). The ragged edges and surfaces of the freshly inflicted wound become

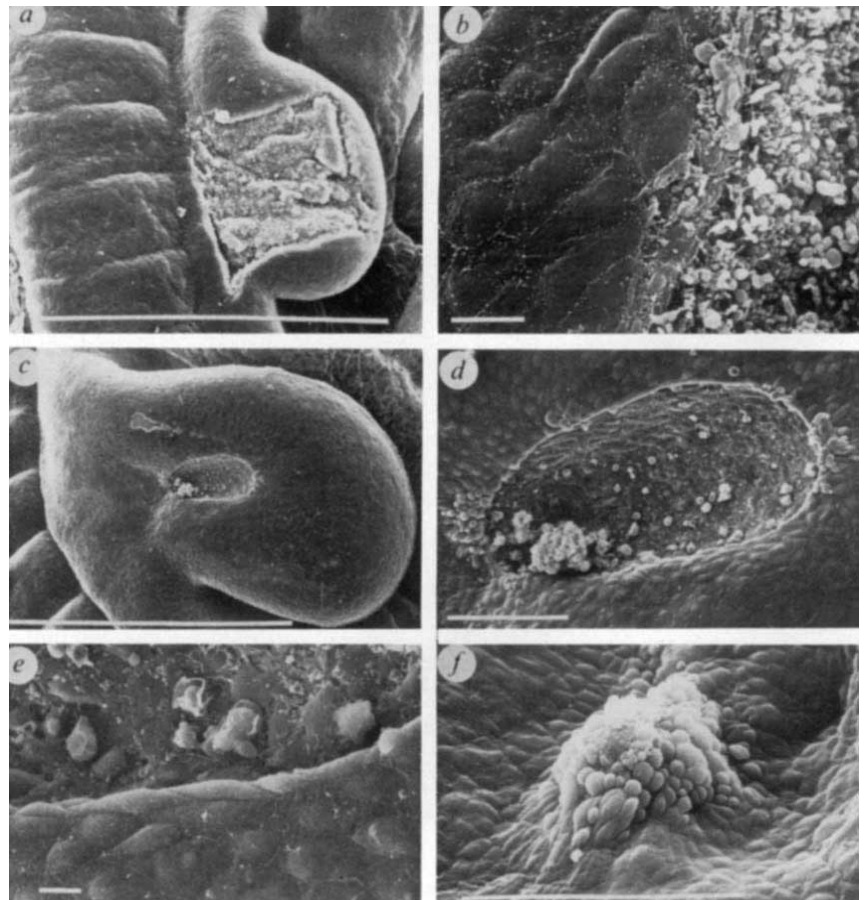
smoothed out within an hour to give a more regular cellular organization that persists until the wound is closed. The exposed mesenchyme flattens to form an almost epithelium-like surface, whereas the free edge of the epidermis forms a smooth arc around most of the perimeter. The epidermal cells at this edge generally appear to adhere tightly to one another, although in a few places they are individually rounded up as though less tightly coherent. But in either case, in studies using light microscopy, scanning electron microscopy (which reveals surface details), or transmission electron microscopy (which lets us see the basal epidermal layer), we do not find lamellipodia at the epidermal wound front in either the basal or the peridermal layer. In cross-section (Fig. 3), the wound front generally has a rounded or only slightly angular profile, rather than the flattened cellular protrusions one might expect at a leading edge that was actively crawling forwards over the exposed mesenchyme^{5,8}. Neither is there any sign that epithelial cells further from the leading edge are putting out lamellipodia or crawling forward as described in some adult systems⁵.

The absence of lamellipodia tallies with another observation: when we grafted a small patch of embryonic skin onto a denuded region of the limb bud surface, we found that the grafted epidermis, far from expanding over the adjacent vacant territory, actually retracted, leaving its own mesenchyme exposed. This behaviour seems at first sight to be the opposite of that seen during wound closure. Both phenomena could, however, be results of the same underlying mechanism: a circumferential tension in the free edge of the epithelium, acting like a purse string, could drive in the one case the contraction of the isolated patch of skin, in the other case the closure of a wound. Such contractile machinery would be expected to involve actin. We therefore examined the distribution of filamentous actin in the healing wounds, by staining with fluorescently labelled phal-

loidin⁹. We kept our specimens as whole mounts, and used a confocal scanning microscope to view thin optical sections in the plane of the wound. We find that a thick band, or cable, of actin runs around almost all of the epidermal wound margin (Fig. 4). This is confined to the free margin of the front row of basal epidermal cells, and appears to be continuous from cell to cell (presumably through adherens junctions¹⁰). As seen in optical section, it has typically about five times the actin content of an ordinary band of cortical actin (Fig. 4i). Transmission electron microscopy of cross-sections of the wound edge confirms the presence of the cable (Fig. 3b). In some regions, the cable is split into several parallel strands, and at a few points around the margin (typically two or three) the continuity of the cable is broken, the actin distribution appears disorderly, and the wound front is irregular rather than smoothly curved; some of these discontinuities may represent sites of stronger attachment to the substratum. The cable is well defined within an hour after wounding (Fig. 4a-c), and other studies (K. Midwinter, J. McCluskey, P. M. and J. L., manuscript in preparation) show that it begins to form within minutes; it then persists until the wound is closed. The thickness of the established cable, after the first hour or two, does not change noticeably as the cable shortens during healing. Moreover, there is no obvious change in the average shape of the marginal basal cells in which the cable runs (Fig. 4g), until the final stages of wound closure (and then only at sites where the wound margin is particularly sharply curved). Cells must therefore be leaving the marginal row as the wound closes (most probably by slipping back into the epithelium to the rear of this row). The pattern of junctional contacts through which the actin cable maintains its appearance of continuity must be continually changing to allow the cable as a whole to shorten. The whole process ends with a 'pile-up', as the epidermal wound front converges at the site of final

FIG. 1 Scanning electron micrographs showing typical stages in the healing of a wound made by removing a square patch of embryonic skin from the dorsal surface of a chick wing bud at 4 days of incubation (stage 22/23). *a*, A wing bud immediately after wounding. Scale bar, 1 mm. *b*, Detail from *a*, showing part of the edge of the fresh wound. Scale bar, 10 μ m. *c*, A wing bud 12 h after wounding. Scale bar, 1 mm. *d*, Detail from *c*, showing the taut, smooth appearance of the free edge of the epidermis at the wound margin and the flattened, epithelioid surface (littered with a certain amount of debris) formed by the exposed mesenchyme. Scale bar, 100 μ m. *e*, Detail from *d*, showing the appearance of the cells at the wound margin; mesenchyme above, epidermis below. Scale bar, 10 μ m. *f*, A recently closed wound, 24 h after wounding. Epidermal cells have been thrown up into a small mound at the site of closure. Scale bar, 100 μ m.

METHODS. The patch of skin was removed with a sharpened tungsten needle, a few drops of Hanks' balanced salt solution were added to keep the wound from drying out, and the egg was then returned to the incubator. Specimens were fixed in half-strength Karnovsky fixative (2% formaldehyde plus 2.5% glutaraldehyde in 0.075 M cacodylate buffer, pH 7.4) at 0, 1, 3, 6, 12, 18 and 24 h after wounding, postfixed in osmium tetroxide, and prepared for scanning electron microscopy in the standard way. At least 3 specimens were examined for each time point. In all cases scale bars show the sizes of the fixed specimens; living specimens, before fixation shrinkage, are $\sim$ 30% larger in linear dimension.



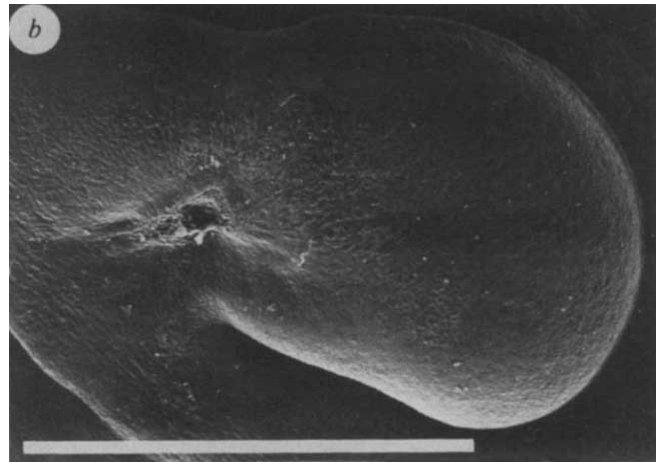


FIG. 2 Evidence that the wound closes by a movement of the epidermis over the mesenchyme. *a*, A wing bud 18 h after wounding. The boundary of the patch of mesenchyme initially denuded of epidermis has been marked with dots of Dil, applied immediately after the operation; these are still visible as red spots 18 h later. *b*, The same wing bud fixed and viewed by scanning electron microscopy to reveal the outlines of the healing epidermis. The free edge of the epidermis has clearly moved in over the

marked wound mesenchyme, leaving less than 10% of it exposed by this stage. Scale bar, 1 mm.

METHODS. Dil was dissolved at a concentration of 1% in ethanol and pressure-ejected onto the exposed mesenchyme, at a series of points around the wound margin, from a glass micropipette with a tip diameter of about 10 μm . Specimens were viewed fresh, by bright-field and epifluorescence optics, before fixation for scanning electron microscopy (see Fig. 1, legend).

closure. Instead of stopping immediately, the motile machinery apparently overruns, throwing up a small volcano-shaped mound of epidermal cells (Fig. 1*e*). Although the cable has by this stage disappeared, the cells in the mound still contain a great deal of filamentous actin in a disorganized arrangement (Fig. 4*h*). All of these phenomena, including the formation of an actin cable, are seen also in the healing of analogous types of wounds in the mouse embryo (P.M., J. McCluskey and J. L., manuscript in preparation).

We have as yet no direct proof that the actin cable provides

the driving force for epidermal wound closure in our system, but this interpretation seems far more plausible than alternatives such as pushing from the rear or hidden crawling movements of cells behind the wound front¹¹. Indeed, our transmission electron micrographs (Fig. 3) show no lamellipodia or other obvious motile structures extending from the epidermal cells behind the wound front. A purse-string mechanism based on a contractile cable fits the observations of others who have investigated embryonic wound healing and seen similar patterns of epidermal movement¹²⁻¹⁴. But it does not apply to minute

FIG. 3 *a*, Transmission electron micrograph of a cross-section through the epidermis at the edge of a wound that has had 12 h to heal. Note the rounded contour of the leading edge and the absence of any flattened lamellipodia extending over the substratum, either from the cells at the leading edge or from those further back. The arrow marks the site where the basal lamina ends; to the rear of this, a normal basal lamina is clearly visible at higher magnification. Scale bar, 10 μm . *b*, Detail of basal epidermal cells at a similar wound edge in a specimen fixed to preserve specifically actin filaments. The arrow indicates the actin cable, seen in cross-section. Scale bar, 2 μm .

METHODS. *a*, Specimens were fixed in half-strength Karnovsky fixative, postfixed in osmium tetroxide, embedded in Araldite, thin sectioned and stained with uranyl acetate and lead citrate in the usual way. This standard fixation procedure does not preserve actin filaments²⁷. *b*, To preserve actin filaments (at the expense of good membrane preservation), we used a protocol from K. McDonald²⁸: limb buds were fixed in cacodylate-buffered 1% glutaraldehyde, rinsed in buffer, treated with 0.8% $\text{K}_3\text{Fe}(\text{CN})_6$ in buffer for 30 min, postfixed with 0.5% osmium tetroxide in buffer for 1 min, embedded in Araldite, sectioned and stained with uranyl acetate and lead citrate.

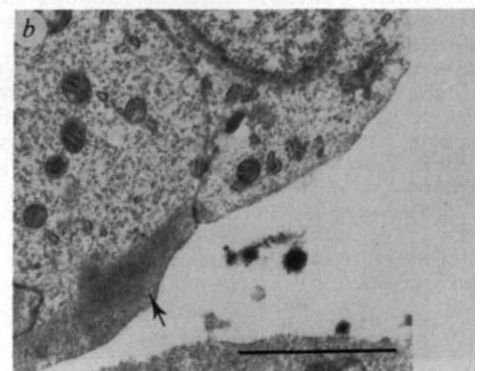
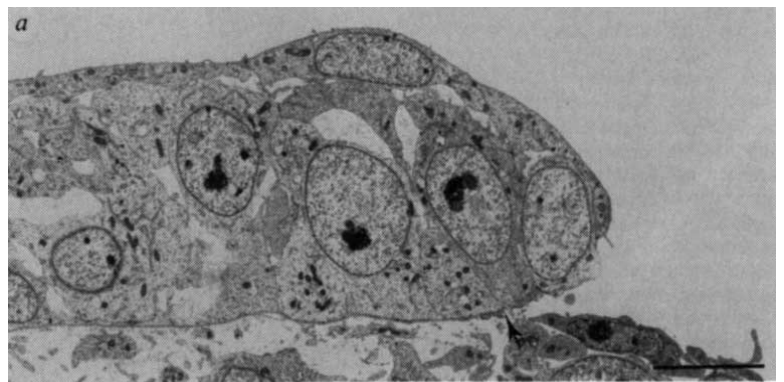
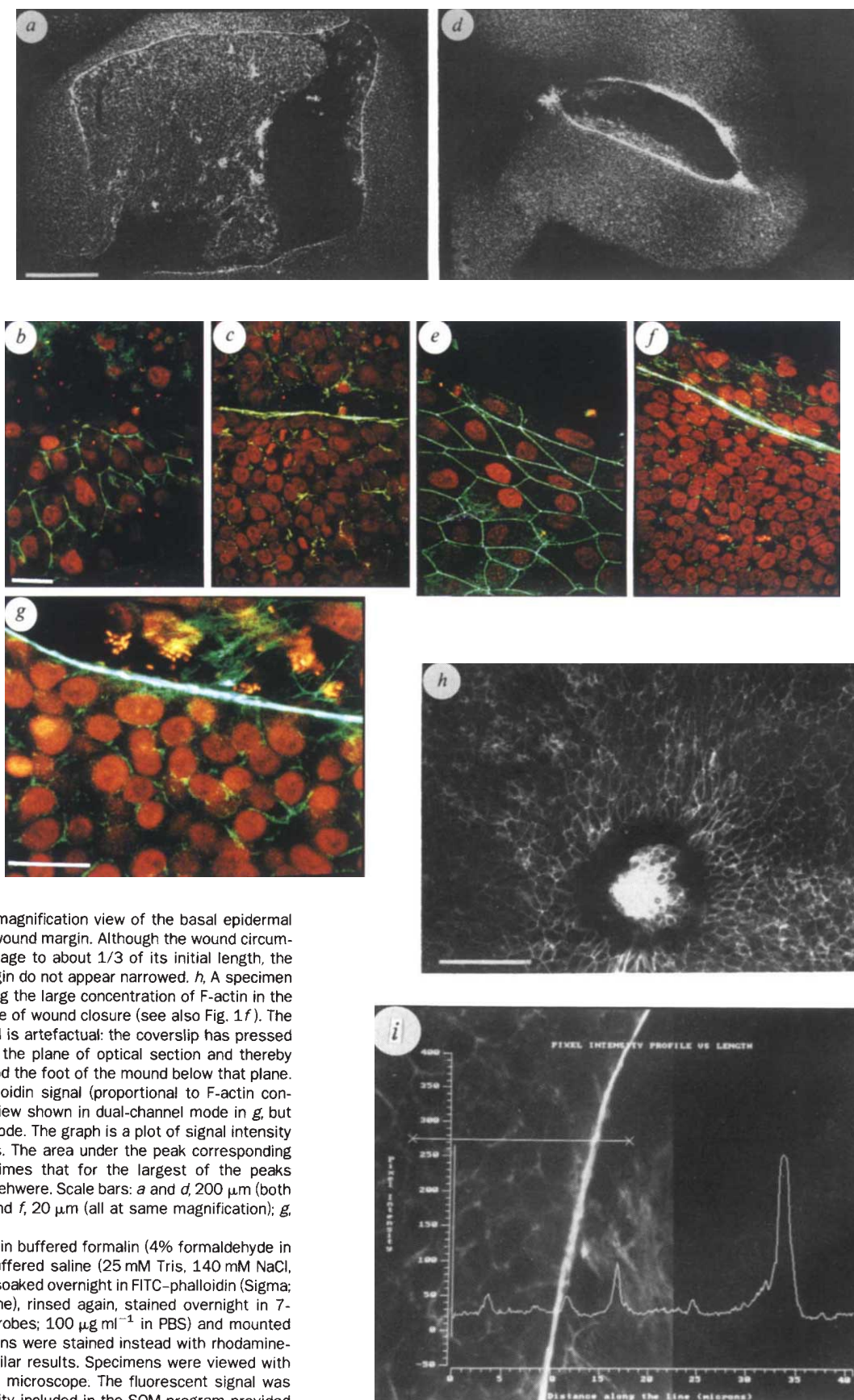


FIG. 4 Optical sections of whole-mount specimens, stained with fluorescein isothiocyanate (FITC)-phalloidin and 7-aminoactinomycin D to reveal filamentous actin (green) and cell nuclei (red). The specimens are lightly flattened under a coverslip, bringing the whole wound area, including the exposed mesenchyme surface, into roughly the same focal plane. *a*, Low magnification view of the whole wound 1 h after wounding; only the actin fluorescence is shown in this photo. Note that the corners of the wound have already become rounded, and that there is already an actin cable at the free edge of the epidermis. *b*, Detail of a part of the wound margin: a section in the plane of the periderm, whose large flat cells are outlined by their cortical actin. *c*, Same part of wound margin as in *b*, but a few μm deeper, in the plane of the basal epidermal cells, showing an actin cable at the wound margin, apparently continuous from cell to cell. Some mesenchyme cells are also visible. *d*, Low magnification view of the whole wound 12 h after wounding. *e*, Detail of the periderm at the wound margin, from the same 12-hour specimen. *f*, Same part of the wound margin as in *e*, but a few μm deeper, in the plane of the basal epidermal cells, showing the actin cable. *g*, High magnification view of the basal epidermal layer of another sample of 12-h wound margin. Although the wound circumference has shortened by this stage to about 1/3 of its initial length, the individual cells at the wound margin do not appear narrowed. *h*, A specimen fixed 24 h after wounding, showing the large concentration of F-actin in the epidermal cells piled up at the site of wound closure (see also Fig. 1*f*). The dark rim around the bright mound is artefactual: the coverslip has pressed the top of the mound down into the plane of optical section and thereby pushed down the epidermis around the foot of the mound below that plane. *i*, Measurement of the FITC-phalloidin signal (proportional to F-actin concentration) in the same field of view shown in dual-channel mode in *g*, but scanned here in single-channel mode. The graph is a plot of signal intensity along the line joining the crosses. The area under the peak corresponding to the actin cable is about 5 times that for the largest of the peaks corresponding to cortical actin elsewhere. Scale bars: *a* and *d*, 200 μm (both at same magnification); *b*, *c*, *e* and *f*, 20 μm (all at same magnification); *g*, 20 μm ; *h*, 100 μm .

METHODS. Specimens were fixed in buffered formalin (4% formaldehyde in PBS) for 1 hour, rinsed in Tris-buffered saline (25 mM Tris, 140 mM NaCl, 3 mM KCl, pH 7.4) for 3×20 min, soaked overnight in FITC-phalloidin (Sigma; $2.5 \mu\text{g ml}^{-1}$ in Tris-buffered saline), rinsed again, stained overnight in 7-aminoactinomycin D (Molecular Probes; $100 \mu\text{g ml}^{-1}$ in PBS) and mounted in buffered saline. Some specimens were stained instead with rhodamine-labelled phalloidin alone, with similar results. Specimens were viewed with a BioRad confocal laser scanning microscope. The fluorescent signal was measured in *i* using the LEN facility included in the SOM program provided with the confocal microscope.



epidermal lesions that leave the basal lamina intact¹⁵, or to embryonic epidermal cells moving on a rigid adhesive substratum in culture^{11,16}; here the conventional mechanism of epidermal crawling seems to operate.

Two central questions arise: what initiates the formation of the actin cable, and what activates its contraction and the associated cell rearrangements? The experiments of Kolega¹⁷ are illuminating here. Stress applied to a migrating epidermal cell at right angles to its direction of movement (tangential to the leading edge of its lamellipodium) causes the lamellipodium to retract and brings about a reorganization of the cell's actin into a cable oriented along the major axis of stress. In our system, we upset the pattern of tension in the epidermis by cutting so as to create a free epidermal edge. This abolishes the component of tension acting at right angles to the cut, while the tangential tension remains, suppressing lamellipodia and provoking formation of an actin cable, which in turn presumably reinforces the tangential tension¹⁸. The act of wounding may also trigger an

influx of Ca^{2+} , activating contraction of the cable¹⁹; and we have in fact shown elsewhere that *c-fos* is rapidly induced in the epidermis at a wound margin in rat embryos²⁰. We do not know how this activated state is subsequently maintained throughout the healing process, although the final overrun and pile-up suggest that it takes some time to switch off. One attractive hypothesis would be that there is positive feedback, whereby tension in the cable generates an intracellular chemical signal for the maintenance of tension; this would maintain an activated state in every cell containing a segment of the taut cable. There is indeed direct evidence that mechanical stress can generate intracellular signals that regulate gene expression^{21,22}.

The actin cable provides a novel way for epidermal cells to coordinate their activities so as to close gaps, fulfilling their function as a self-repairing barrier layer. It will be interesting to see whether similar actin cables also play a part in driving the coordinated movements of normal morphogenesis²³⁻²⁶. □

Received 31 July; accepted 24 September 1992.

1. Odland, G. & Ross, R. *J. Cell Biol.* **39**, 135-151 (1968).
2. Krawczyk, W. S. *J. Cell Biol.* **49**, 247-263 (1971).
3. Pang, S. C., Daniels, W. H. & Buck, R. C. *Am. J. Anat.* **153**, 177-192 (1978).
4. Bereiter-Hahn, J. in *Biology of the Integument* (eds Bereiter-Hahn, J., Matoltsy, A. G. & Richards, K. S.) 443-471 (Springer, New York, 1986).
5. Buck, R. C. *Invest. Ophthalmol. Vis. Sci.* **18**, 767-784 (1979).
6. Gabbiani, G. *Int. Rev. Cytol.* **48**, 187-219 (1977).
7. Honig, M. G. & Hume, R. I. *Trends Neurosci.* **12**, 333-341 (1989).
8. Trinkaus, J. P. *Cells into Organs: the Forces that Shape the Embryo* (Prentice-Hall, Englewood Cliffs, 1984).
9. Wulf, E., Deboen, A., Bautz, F. A., Faulstich, H. & Wieland, T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4498-4502 (1979).
10. Geiger, B., Volk, T., Volberg, T. & Bendori, R. *J. Cell Sci.* (suppl.) **8**, 251-272 (1987).
11. DiPasquale, A. *Exp. Cell Res.* **94**, 191-215 (1975).
12. Thévenet, A. *Archs Anat. microsc. Morph. exp.* **70**, 227-244 (1981).
13. Stanisstreet, M., Wakely, J. & England, M. A. *J. Embryol. exp. Morphol.* **59**, 341-353 (1980).
14. Smedley, M. J. & Stanisstreet, M. *J. Embryol. exp. Morphol.* **83**, 109-117 (1984).

15. Radice, G. P. *Dev. Biol.* **76**, 26-46 (1980).
16. Vaughan, R. B. & Trinkaus, J. P. *J. Cell Sci.* **1**, 407-413 (1966).
17. Kolega, J. *J. Cell Biol.* **102**, 1400-1411 (1986).
18. Sherratt, J. A. & Lewis, J. *Bull. Math. Biol.* (in the press).
19. Stanisstreet, M. *J. Embryol. exp. Morphol.* **67**, 195-205 (1982).
20. Martin, P. & Nobes, C. D. *Mech. Dev.* **38**, 209-216 (1992).
21. Komuro, I. et al. *J. Biol. Chem.* **266**, 1265-1268 (1991).
22. Ben-Ze'ev, A. *BioEssays* **13**, 207-212 (1991).
23. Fristrom, D. *Tissue Cell* **20**, 645-690 (1988).
24. Bard, J. *Morphogenesis: The Cellular and Molecular Processes of Developmental Anatomy* (Cambridge Univ. Press, Cambridge, 1990).
25. Belousov, L. V., Dorfman, J. G. & Cherdantsev, V. G. *J. Embryol. exp. Morphol.* **34**, 559-574 (1975).
26. Odell, G. M., Oster, G., Alberch, P. & Burnside, B. *Dev. Biol.* **85**, 446-462 (1981).
27. Maupin-Szamer, P. & Pollard, T. D. *J. Cell Biol.* **77**, 837-852 (1978).
28. McDonald, K. *J. ultrastruct. Res.* **86**, 107-118 (1984).

ACKNOWLEDGEMENTS. We thank B. Luke, M. Masih and N. White for their help with EM and confocal scanning microscopy, and G. Dunn, G. Ireland, S. Jesuthasan and C. Stern for comments.

Engineering galactose-binding activity into a C-type mannose-binding protein

Kurt Drickamer

Department of Biochemistry and Molecular Biophysics,
Columbia University, New York, New York 10032, USA

CALCIUM-DEPENDENT or C-type carbohydrate-recognition domains are homologous protein modules found in a variety of animal lectins¹. Selective binding of sugars by these domains is essential for glycoprotein clearance, cell-cell adhesion and pathogen neutralization. Although various C-type carbohydrate-recognition domains share sequence identity ranging from 20 to 55%, their sugar-binding characteristics vary widely. The structure of a mannose-binding carbohydrate-recognition domain in complex with a saccharide ligand suggests that two glutamic acid-asparagine pairs are essential determinants of ligand binding by this domain². In C-type lectins that bind galactose with higher affinity than mannose, one of these pairs is replaced by glutamine-aspartic acid. Here we shift the sequence of the mannose-binding protein to correspond to that found in galactose-binding domains in order to test the importance of these residues in sugar-binding selectivity. This simple switch in the position of a single amide group alters the binding activity of the domain so that galactose becomes the preferred ligand.

The crystal structure reveals that mannose binds to the carbohydrate-recognition domain (CRD) of mannose-binding protein in direct coordination to one of the calcium ions (designated

number 2)². Amino-acid residues that ligate both this calcium ion and the saccharide are located in the C-terminal 25 amino acids of the 115-amino-acid domain (Fig. 1). Residues corresponding to Asp 206 (a calcium ligand), and to Glu 193 and Asn 205 (calcium ligands that form hydrogen bonds to the 4-OH group of mannose) are found in all CRDs known to bind sugars. Residues Glu 185 and Asn 187 are calcium ligands that form hydrogen bonds to the 3-OH of mannose. These residues are conserved in CRDs that bind mannose or glucose, ligands with similarly disposed equatorial 3- and 4-OH groups. But in lectins that bind galactose, glutamine and aspartic acid are consistently found at these positions. Particularly interesting are the CRDs of two group II proteins (Fig. 1), the *N*-acetylglucosamine-binding from hepatic lectin from chicken³ and a placental mannose receptor⁴. Although these CRDs are most closely related in overall sequence to the galactose-binding asialoglycoprotein

TABLE 1 Inhibition of neoglycoprotein binding by monosaccharides

	K_i (mM)		Wild type
	Wild type	QPD mutant	Mutant
Mannose	7.2 ± 0.8	18 ± 3	0.40
Glucose	13 ± 2	20 ± 2	0.65
<i>N</i> -acetylglucosamine	11 ± 2	44 ± 11	0.25
Fucose	5.6 ± 1.1	4.2 ± 0.3	1.3
Galactose	100 ± 8	5.1 ± 0.1	19.6
<i>N</i> -acetylgalactosamine	>1,000	10.4 ± 0.5	>100

Competition assays were done as for Fig. 3, using radioiodinated mannose₂₃-BSA as the reporter ligand for wild-type protein and galactose₃₄-BSA for the QPD mutant.

receptor, they resemble mannose-binding protein in that they too contain glutamic acid and asparagine residues at the 3-OH binding positions in mannose-binding protein. These observations suggest that the acid-amide side-chain arrangement at these two positions may be a primary determinant of ligand-binding specificity in the C-type lectins.

To test the importance of Glu 185 and Asn 187, a mutant was constructed in which these residues are simultaneously replaced with Gln and Asp, to conform to the Gln-Pro-Asp (QPD) galactose-type sequence. This protein was expressed in the bacterial system previously used to express the wild-type CRD⁵. Unlike the wild-type protein, the mutant does not bind significantly to mannose-Sepharose, but it is retarded on galactose-Sepharose. The behaviour of the mutant protein is compared with the wild-type in Fig. 2. Mutant protein retarded on the column was pooled and used for a more detailed analysis of sugar binding.

Binding specificities of wild-type and QPD mutant CRDs were compared using a solid-phase assay, in which radioiodinated neoglycoprotein derivatives of bovine serum albumin (BSA) were bound to CRDs immobilized in polystyrene wells. Mannose-BSA but not galactose-BSA binds to wild-type CRD, but both ligands bind to the mutant CRD. However, binding of galactose-containing ligand is ~10 times higher than the binding of mannose-containing ligand (data not shown). Monosaccharide-binding affinities were determined using these test ligands in a competition assay (Fig. 3). Mannose competes for

binding to wild-type CRD with 14-fold higher affinity than does galactose ($K_{i,Man} = 7.2$ mM versus $K_{i,Gal} = 100$ mM). In contrast, galactose is a better competitor of neoglycoprotein binding to the mutant CRD by more than threefold ($K_{i,Man} = 18.4$ mM versus $K_{i,Gal} = 5.1$ mM). Similar results are obtained for the mutant when either galactose-BSA or mannose-BSA is used as the reporter ligand, indicating that the same binding site is detected in both assays.

To assess which portions of the galactose ligand interact with the QPD mutant, the inhibitory activities of several monosaccharides were compared. Because binding of *N*-acetyl-galactosamine is almost as strong as binding to galactose (Table 1), it seems likely that the shared equatorial-axial arrangement of hydroxyl groups 3 and 4 in these sugars plays a major role in binding. In this respect, the QPD mutant mimics the behaviour of the natural galactose-binding C-type lectins, such as the asialoglycoprotein receptor⁶. The $K_{i,Gal}$ of 5.1 mM also compares favourably with the $K_{i,Gal}$ of 1.7 mM for the asialoglycoprotein receptor⁷. Binding of all ligands that contain equatorial 3 and 4 hydroxyl groups (mannose, glucose, and *N*-acetylglucosamine) is decreased in the mutant.

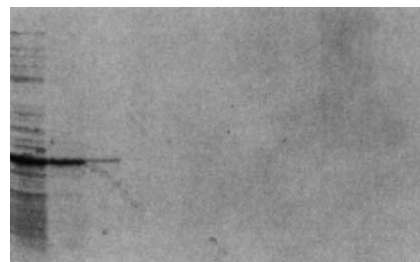
The calcium affinity of the QPD mutant estimated from the calcium dependence of ligand binding in the solid phase assay⁸ is very similar to wild-type ($K_{Ca} = 0.60$ mM for wild-type and 0.73 mM for the mutant; data not shown). This observation is consistent with proteolysis studies, because formation of a protease-resistant conformation⁵ shows similar calcium dependence

			185 187	193	205 206	
			↓ ↓	↓	↓ ↓	
	CARTILAGE PG (Hu)		WRPNQPDNFFAA	---GEDCVMI---	WHEGGEWNDVPCN-YHLPFTC	
	CARTILAGE PG (Ra)		WRPNQPDNFFAT	---GEDCVMI---	WHERGEWNDVPCN-YQLPFTC	
	CARTILAGE PG (Ch)		WRPNQPDNFFAA	---GEDCVMI---	WHEGGEWNDVPCN-YHLPFTC	
	FIBROBLAST PG (Hu)		WRPNQPDNFFSA	---GEDCVMI---	WHENGWNDVPCN-YHLYTTC	
			↓ ↓	↓	↓ ↓	
Gal	Mφ Gal-REC (Ra)		WAPKQPDNWHGHLGGEDCAHFT	---SDGRWNDVQC-RPYRWVC		
	HEP LEC-1 (Ra)		WRPGQPDNWHGHLGGEDCAHFT	---TDGRWNDVQC-RPYRWVC		
	HEP LEC-2 (Hu)		WRPEQPDNWHGHLGGEDCAHFT	---DDGRWNDVQC-RPYRWVC		
	HEP LEC-2/3 (Ra)		WAFQPDNWHGHEGGEDCAEIL	---SDGLWNDVFC-QVNRWAC		
	HEP LEC-2 (Hu)		WAVQPDNWHGHEGGEDCCEVQ	---PDGRWNDVFC-QVNRWVC		
	KUP Fuc-REC (Ra)		WRKQPDNWR-HGNGEREDCVHLQ	---RMWNDMACG-TAYNRVC		
			↓ ↓	↓	↓ ↓	
	IgE Fc REC (Mo)		WNPGEPNNGGQ	---GEDCVMMR---	GSQWNDVFCRSLDAMVC	
	IgE Fc REC (Hu)		WAPGEPTSRSQ	---GEDCVMMR---	GSQWNDVFCRDLKGAWVC	
			↓ ↓	↓	↓ ↓	
	PLC Man-REC (Hu)		WNRGEPNNVGE	---EDCAEFS---	GNQWNDVCKN---LAKWIC	
	HEP LEC (Ch)		WKEGEPNNRGF	---NEDCAHVW---	TSGQWNDVYCT-YECYVVC	
			↓ ↓	↓	↓ ↓	
Man	MBP-A (Ra)		WKKDEPNHGS	---GEDCVTIIV---	DNLGWNDSQC-ASHTAVC	
	MBP (Hu)		WNEGEPNNAGS	---DEDCVLLL---	KNGQWNDVPCS-TSHLAVC	
	MBP-C (Ra)		WNEGEPNNVGS	---GENCVLVK---	TNGQWNDVPCS-DSFLVVC	
	CONGLUTININ (Bo)		WADGEPNNSE	---GQPCNCEIF---	PDGQWNDVPCS-KQLLVIC	
	SP-D (Hu)		WAPGEPNDGG	---SEDCVEIF---	TNGQWNDVPCS-EKRLVVC	
			↓ ↓	↓	↓ ↓	
	SP-A (Do)		WYPGEPRGRGK	---EQCVEMY---	TDGQWNNKNC-QYRLAIC	
	SP-A (Hu)		WYRGEPRGRGK	---EQCVEMY---	TDGQWNNKNC-YSRLTIC	
			↓ ↓	↓	↓ ↓	
	L-SELECTIN (Mo)		WGAGEPNKKKS	---KEDCVEIYIKRERDSGKWNDDACH-KRKAALC		
	L-SELECTIN (Hu)		WGDGEPNNKKK	---KEDCVEIYIKRNDAGKWNDDACH-KLKAALC		
	E-SELECTIN (Hu)		WAPGEPNNRQK	---DEDCVEIYIKREKDVGMWNDERCS-KKLLALC		
	P-SELECTIN (Hu)		WADNEPNKKRN	---NEDCVEIYIKSPAGKWNDEHCL-KKKHALC		
			↓ ↓	↓	↓ ↓	
	Mφ ManR CRD-4 (Hu)		WAYGEPNNYQN	---VEYCGELK---	GDPTMSWINDINCE-HLNNWIC	
	Mφ ManR CRD-5 (Hu)		WATGEPNFANE	---DENCVTMY---	SNSGFWDINDINC-YPNAFIC	
	Mφ ManR CRD-2 (Hu)		WLRGEPSHENN	---RQEDCVVMK---	GKDGWADRGCE-WPLGYIC	
	Mφ ManR CRD-7 (Hu)		WAADEPKLSKA	---CVYLD---	LDGYWKAHCN-ESFYFLC	
	Mφ ManR CRD-8 (Hu)		WNTGDPSGERN	---DCVALH---	ASSGFWSNIHCS-SYGYIC	
	Mφ ManR CRD-1 (Hu)		WLPGSPSAEPG	---KSCVSLN---	PGNAKWENLEVC-QKLGIC	
	Mφ ManR CRD-3 (Hu)		WNSDMPRKPG	---CVAMR---	TGAGGLWDLVLC-EKAKFVC	
	Mφ ManR CRD-6 (Hu)		WGKGYPGRRSSL	SYEDACVVI	IGGSNAEAGKWMDDTCD-SKRGYIC	

FIG. 1 Alignment of carboxy-terminal portions of C-type CRD amino-acid sequences. CRDs in groups I to IV and VI have been identified as evolutionarily related on the basis of overall similarity of the CRDs¹³. An approximate categorization of monosaccharide-binding activity is presented on the left. Arrows indicate five ligands for calcium ion 2 seen in the crystal structure². PG, proteoglycan; REC, receptor; LEC, lectin; MBP, mannose-binding protein; SP, pulmonary surfactant apoprotein; Mφ, macrophage; HEP, hepatocyte; KUP, Kupffer cell; PLC, placenta; Hu, human; Ra, rat; Ca, chicken; Mo, murine; Bo, bovine; Do, dog. Citations for sequences may be found in refs 4, 13 and 14.

FIG. 2 Preparation of QPD mutant of mannose-binding protein CRD. The expression plasmid for the CRD⁵ was altered by replacement of an *AlwNI* to *BstXI* fragment with a synthetic double-stranded oligonucleotide (5'-CTGGAAAAAGGATCAGCCCGATGACCATGGCT-3' annealed with 5'-ATGGCATCGGGCTGATCCTTTTCCAGTTG-3'), corresponding to bases 546-580 in the original cDNA sequence¹⁵. Recombinant DNA procedures used standard protocols¹⁶. Wild-type and mutant CRDs were expressed as previously described⁵, except that guanidine-solubilized protein was not diluted before dialysis. After renaturation, soluble material was applied to a 1 × 11 cm (10 ml) column of galactose-Sepharose¹⁷, which was then rinsed with 20 ml calcium-containing loading buffer and 20 ml EDTA-containing eluting buffer⁵. Fractions of ~2 ml were collected during the wash (fractions 1-8) and elution (fractions 9-18) steps, and 25- μ l aliquots of fractions 4 to 14 were analysed on 17.5% SDS-polyacrylamide gels¹⁸ with Coomassie blue. Purified mutant protein in fractions 6 to 8 was used for solid-phase binding assay.

Wild type
4 5 6 7 8 9 10 11 12 13 14



QPD mutant

4 5 6 7 8 9 10 11 12 13 14

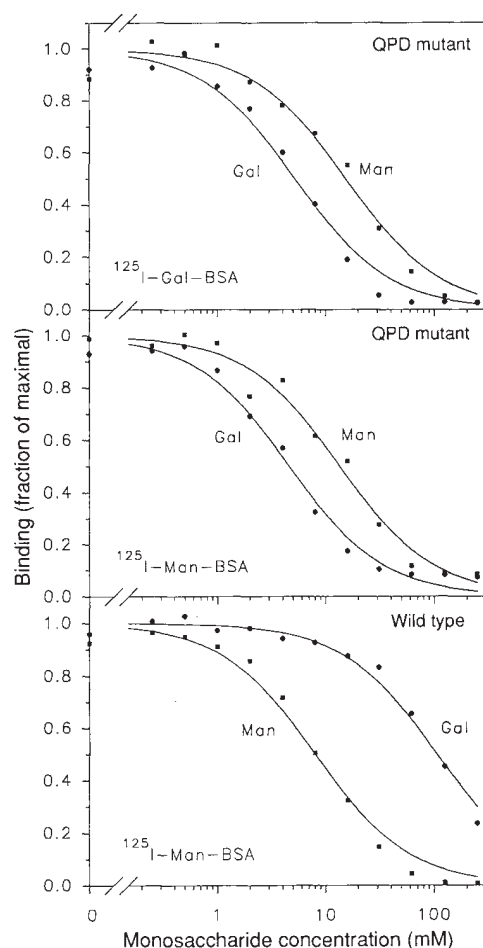
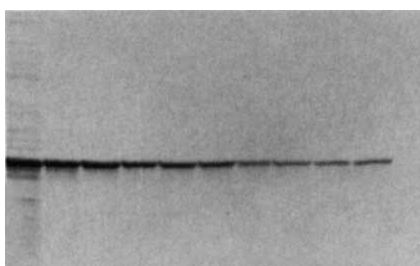


FIG. 3 Monosaccharide competition for binding of radioiodinated neoglycoprotein ligands to immobilized wild-type and mutant CRDs. Binding assay procedures have been described⁶. Galactose₃₄-BSA and mannose₂₃-BSA (E-Y Lab, San Mateo, CA) were used as test ligands. Each point was determined in duplicate. K_i values (concentration of monosaccharide giving 50% inhibition of neoglycoprotein binding) reported in the text and in Table 1 are averages of three independent experiments.

(K_{Ca} = 0.72 mM for wild type and 0.39 mM for the mutant; data not shown). Thus, as expected from the structure of the carbohydrate-recognition domain², modification of the amide position does not appear to affect ligation of calcium ion 2.

In the light of these results, it is interesting to examine the pattern of ligand-binding activity in other C-type CRDs shown in Fig. 1. Two CRDs in the macrophage mannose receptor, numbers 4 and 5, contain the characteristic mannose-binding residues. These two domains are essential for high-affinity binding of mannose-containing ligands by this receptor⁹. Proteins that do not contain either the Glu-Pro-Asn sequence of mannose/glucose-binding CRDs or the Gln-Pro-Asp sequence of galactose-binding CRDs have unique ligand-binding characteristics. For example, pulmonary surfactant apoprotein SP-A, which contains the sequence Glu-Pro-Ala, binds a variety of sugars with weak affinity, including both mannose and galactose^{10,11}. Finally, the IgE Fc receptor of human lymphocytes (CD23) binds a protein rather than a carbohydrate ligand¹², and also does not contain either the galactose or mannose-characteristic sequences.

The present findings are consistent with the crystallographic results showing that Glu 185 and Asn 187 in the sequence Glu-Pro-Asn are major determinants of mannose-binding activity, and provide strong evidence that the Gln-Pro-Asp sequence is probably important in galactose binding by other C-type CRDs. Comparative ligand-binding studies suggest that 3- and 4-hydroxyls are important for ligand binding by all of the C-type animal lectins⁶. It will be of interest to see how these groups in an equatorial-axial arrangement in galactose interact with the Gln-Pro-Asp sequence.

From the structure of the CRD, it is not obvious how the subtle change in the QPD mutant reduces the affinity of the domain for mannose, because modelling suggests that a pair of equatorial hydroxyl groups would still be accommodated even with the modified sequence. Similarly, it is unclear from the sequences how the natural galactose-binding CRDs discriminate more effectively against mannose than does the QPD mutant. Partial loss of affinity for mannose and dramatically increased affinity for galactose and *N*-acetylgalactosamine may result from an altered conformation of the CRD in the region near the ligand-binding site. This conformation would probably not be

as stable in the mutant as it is in the natural galactose-binding CRDs. It should also be noted that many of the proteins within the broad categories of mannose/glucose-binding and galactose-binding in Fig. 1 bind to distinct subsets of sugars. Thus there must be additional structural features of the CRDs that influence their interactions with ligands. □

Received 31 July; accepted 28 September 1992.

- Drickamer, K. *J. biol. Chem.* **263**, 9557–9560 (1988).
- Weis, W. I., Drickamer, K. & Hendrickson, W. A. *Nature* **360**, 127–134 (1992).
- Sikder, S. K., Kabat, E. A., Steer, C. J. & Ashwell, G. *J. biol. Chem.* **258**, 12520–12525 (1983).
- Curtis, B. M., Scharnowski, S. & Watson, A. J. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8356–8360 (1992).
- Weis, W. I., Crichtlow, G. V., Murthy, H. M. K., Hendrickson, W. A. & Drickamer, K. *J. biol. Chem.* **266**, 20678–20686 (1991).
- Lee, R. T. *et al. J. biol. Chem.* **266**, 4810–4815 (1991).
- Connolly, D. T., Townsend, R. R., Kawaguchi, K., Bell, W. R. & Lee, Y. C. *J. biol. Chem.* **257**, 939–945 (1982).
- Quesenberry, M. S. & Drickamer, K. *J. biol. Chem.* **267**, 10831–10841 (1992).
- Taylor, M. E., Bezouska, K. & Drickamer, K. *J. biol. Chem.* **267**, 1719–1726 (1992).
- Haagsman, H. P. *et al. J. biol. Chem.* **262**, 13877–13880 (1987).
- Childs, R. A. *et al. J. biol. Chem.* **267**, 9972–9979 (1992).
- Vercellij, D. *et al. Nature* **338**, 649–651 (1989).
- Bezouska, K., Crichtlow, G. V., Rose, J. M., Taylor, M. E. & Drickamer, K. *J. biol. Chem.* **266**, 11604–11609 (1991).
- Taylor, M. E., Conary, J. T., Lennartz, M. R., Stahl, P. D. & Drickamer, K. *J. biol. Chem.* **265**, 12156–12162 (1990).
- Drickamer, K., Dordal, M. S. & Reynolds, L. J. *J. biol. Chem.* **261**, 6878–6887 (1986).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1982).
- Fornstedt, N. & Porath, J. *FEBS Lett.* **57**, 187–191 (1975).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).

ACKNOWLEDGEMENTS. I thank N. Ruiz for technical assistance, and W. Weis and W. Hendrickson for discussion and critical comments on the manuscript. This work was supported by a grant from the National Institutes of Health. K.D. is recipient of a Faculty Salary Award from the American Cancer Society.

Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography

Zsolt Böcskei*, Colin R. Groom, Darren R. Flower*, Charles E. Wright, Simon E. V. Phillips, Andrea Cavaggioni†, John B. C. Findlay & Anthony C. T. North‡

Department of Biochemistry and Molecular Biology, The University of Leeds, Leeds LS2 9JT, UK

† Istituto di Fisiologia Umana, Università di Parma, 43100 Parma, Italy

THE principal protein excreted in male rat urine, urinary α 2-globulin and the homologous mouse protein, major urinary protein, have been well characterized, although their functions remain unclear. Male rat urine affects the behaviour and sexual response of female rats¹, leading to the proposal that rodent urinary proteins are responsible for binding pheromones and their subsequent release from drying urine². Urinary α 2-globulin is also involved in hyaline droplet nephropathy, an important toxicological syndrome in male rats resulting from exposure to a number of industrial chemicals and characterized by the accumulation of liganded urinary α 2-globulin in lysosomes in the kidney, followed by the induction of renal cancer³. We now report the three-dimensional structures of mouse major urinary protein (at 2.4 Å resolution) and rat urinary α 2-globulin (at 2.8 Å resolution). The results corroborate the role of these proteins in pheromone transport and elaborate the structural basis of ligand binding.

The sequences of rat urinary α 2-globulin (A2U) and mouse

TABLE 1 Data collection and structure solution

Crystal	Resolution (Å)	Completeness*	R_{sym} † (%)	Δ_{iso} ‡ (%)	Number of sites	R_{centric} § (%)
MUP native	14–2.4	85	3.3			
HgBr ₂	14–2.5	91	3.1	26	1	49 (3 Å)
K ₂ PtCl ₆	14–3.0	91	5.2	19	5	54 (4 Å)
K ₃ UO ₂ F ₅	14–3.0	95	3.6	14	3	54 (5 Å)
A2U native	14–2.8	97	5.5			

Crystals of MUP were grown from a solution of CdCl₂ in malate buffer (P4₃2₁2, $a=b=57.3$ Å, $c=109.0$ Å), and crystals of A2U were grown from a solution of polyethylene glycol 3350 in citrate buffer (P2₁, $a=56.6$ Å, $b=103.5$ Å, $c=62.7$ Å, $\beta=95.3^\circ$)⁷. No other external ligands were added. Diffraction data were obtained from these crystals on a Nicolet/Siemens X100A area detector, mounted on a Rigaku RU-200 rotating anode X-ray generator. The frame data were processed with the program XDS¹¹. Strong anomalous scattering from Cd²⁺ ions present in the native crystals of MUP prevented the use of anomalous data for initial phase determination. With an overall figure of merit for the MIR phases of only 0.48 at 3 Å resolution, automatic solvent flattening¹² failed to produce sensible molecular boundaries and we therefore defined the molecular envelope manually using the program MEDIT¹³. Solvent flattening then produced an interpretable map. X-PLOR¹⁴ and Hendrickson-Konnert least-squares restrained refinement¹⁵ has led to a final R -factor of 18.7% at 2.4 Å, with an r.m.s. bond-length deviation of 0.021 Å. The present model comprises residues 5–161 (of which the side chains of four residues are disordered), two fully and two partially occupied Cd²⁺ ions, 70 water molecules and a bound molecule of 2-(*sec*-butyl)thiazoline. With this MUP structure, with ligand, water and side-chain atoms uncommon to A2U removed as the basis for a search model, the program MERLOT¹⁶ gave clear, unambiguous solutions for the orientations and positions of the four monomers per asymmetric unit of A2U. The resulting A2U model was then subjected to rigid body minimization by use of X-PLOR, limited rebuilding by FRODO¹⁷, simulated annealing refinement with X-PLOR and conventional restrained least-squares refinement by PROLSQ, to an R factor of 23.4% at 2.8 Å, with an r.m.s. bond-length deviation of 0.023 Å. This model comprises residues 5–161 of four independent protein monomers (of which 3, 2, 3 and 6 residues in the four chains respectively cannot yet be modelled). Solvent molecules have not yet been included.

* Per cent of total possible.

† $R_{\text{sym}} = \sum_n \sum_i |I(h_i) - \langle I(h) \rangle| / \sum_n \sum_i I(h_i)$, where h are unique reflection indices.

‡ Fractional isomorphous difference.

§ $R_{\text{centric}} = \sum_n (|F_{\text{PH}} - F_{\text{P}}| - |F_{\text{H}}|) / \sum_n |F_{\text{PH}} - F_{\text{P}}|$.

major urinary protein (MUP) show them both to be members of the lipocalin protein family². These are small proteins capable of binding hydrophobic molecules with high affinity and selectivity. This burgeoning family now contains over 20 different proteins, principally identified through sequence homology⁴. An alignment showing several lipocalins is given in Fig. 1. Lipocalins have a wide range of functions and include odour-binding proteins found in the nasal mucosa⁵ or excreted in urine. Some members of this family may possess dual molecular recognition properties—for the ligand bound by the protein and for a receptor cell to which the ligand is transported. This has been found for retinol-binding protein⁶ and the odour-binding proteins (M. Boudjelal, A. Sivaprasadarao and J.B.C.F., unpublished results).

We have previously reported the crystallization of MUP and A2U (ref. 7). Although sequence analysis shows them to be lipocalins, molecular replacement using other lipocalin structures as search models failed to produce interpretable solutions; we therefore used the multiple isomorphous replacement method and prepared three heavy-atom derivatives of MUP. The MUP structure was then used as a model to solve A2U by molecular replacement. The crystallographic procedures, summarized in Table 1, will be reported in detail elsewhere.

The structures of MUP and A2U are similar to those of other lipocalins, consisting of an eight-stranded β -barrel and an α -helix, with loops between the β -strands. There are no large differences in main-chain conformation between the four molecules of A2U in the asymmetric unit (r.m.s. difference between the α -carbon (C_α) coordinates for any pair of monomers is 0.76 to 1.20 Å) or between any of these molecules and MUP (r.m.s. difference between all C_α coordinates is 0.96 to 1.20 Å). Their similarity to other lipocalins is typified by an r.m.s. difference in C_α positions of 1.21 Å between equivalent residues in the β -sheet cores of MUP and RBP. The packing

* Present addresses: Chinoin Pharmaceuticals, PO Box 110, H1325 Budapest, Hungary (Z.B.); Department of Physical Chemistry, Fisons Pharmaceuticals, Bakewell Road, Loughborough, Leicestershire LE11 0RH, UK (D.R.F.).

‡ To whom correspondence should be addressed.



FIG. 1 Sequence alignment, made in the light of structural correspondence, of human serum retinol-binding protein (RBP)¹⁸, bovine β -lactoglobulin (BLG)¹⁹, MUP^{20,21} and A2U²². Although only one sequence for MUP and A2U is shown the urinary proteins are encoded by a gene family expressing at least seven closely related forms of the protein. The alignment was produced with the program SOMAP²³. Colours represent the nature of the residues: red, negatively charged; blue, positively charged; green, polar; grey, hydrophobic; purple, aromatic; brown, glycine and proline; yellow, cysteine. Despite

little sequence similarity (typically 30% identity between pairs), the lipocalin family shows a high degree of structural similarity²⁴. RBP, BLG (which also binds retinol and other hydrophobic ligands), insecticynin²⁵ and bilin-binding protein²⁶, the first examples of known 3-dimensional structure, all comprise similar cup-shaped antiparallel β -barrel frameworks with successive +1 topology, formed from strands linked by loops that differ between the proteins.

of MUP and A2U in the crystals is, however, different (Fig. 2) and comparison with the packing observed in other lipocalin structures suggests that there is no consistent pattern in their quaternary structures.

The interiors of the β -barrels of MUP and A2U form pockets with a highly apolar lining, ideal for the transportation of small hydrophobic molecules through hydrophilic media. There is

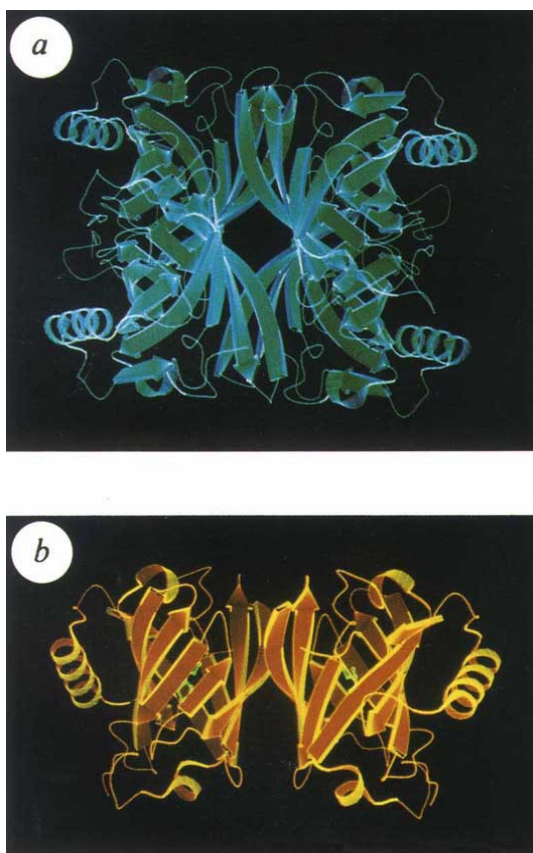


FIG. 2 Ribbon diagram²⁷ of *a*, the A2U tetramer, and *b*, the MUP dimer. A2U crystallizes as a tetramer, with the sides of the β -barrels packing against each other. The one MUP molecule in the asymmetric unit makes its most extensive interactions to a molecule related to it by a 2-fold rotation about an axis parallel to the axis of the β -barrel. Two of the four Cd^{2+} ions in the asymmetric unit, which are necessary for crystallization, are involved in this interface suggesting the stability of crystal packing in MUP is dominated by the interactions with the metal ions.

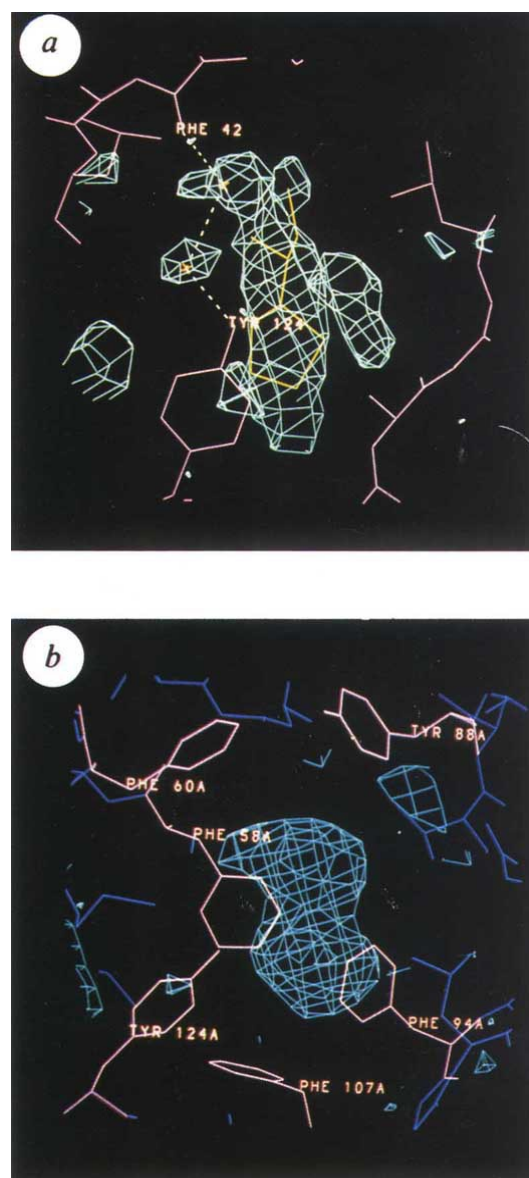


FIG. 3 The electron density in the binding pocket of *a*, MUP with 2-(sec-butyl)thiazoline shown in yellow, and *b*, A2U. Aromatic residues lining the binding pocket are shown in purple. Both are $F_o - F_c$ electron-density maps; the bound ligand was not included in either the refinement or map calculation.

clear electron density in the ligand-binding pocket of MUP and A2U, showing that both have been co-purified and crystallized with a bound ligand. Gas chromatography of MUP suggests that the principal component bound by MUP *in vivo* is 2-(*sec*-butyl)thiazoline⁸ (Fig. 3), although this may represent the major species of several present in the crystal, MUP has been shown to bind at least three different ligands *in vivo* and some or all of these may also be present at low occupancy. The fit to the electron density of 2-(*sec*-butyl)thiazoline is good, but difference electron density maps show additional weak positive electron density close to the modelled ligand. The electron density observed in the ligand-binding pocket of A2U is different in shape, located closer to one side of the barrel and slightly less deep in the binding pocket than in MUP (Fig. 3), although it is still consistent with the size and shape of a number of pheromones.

There are differences in both the identity and positions of amino-acid residues forming the binding pocket of MUP and A2U. Most noticeable are the substitution of Val 58 and Ala 107 of MUP by phenylalanine in A2U. With Phe 94 these three residues form the base of the binding pocket of A2U, forcing any bound ligand to occupy a slightly higher position in the pocket. The substitution of Leu 44 and Ala 122 of MUP by valine in A2U and large movements of the side chains of Phe 60 and Tyr 121 also give rise to important differences in the two binding pockets. Also of note in the MUP structure is a hydrogen bond from the side-chain hydroxyl group of Tyr 124 to one of two water molecules in the binding pocket. In A2U the hydroxyl group of Tyr 124 points directly at the electron density corresponding to bound ligand. These differences help explain why these homologous proteins bind a variety of ligands with different affinities⁹. The size and hydrophobicity of the binding pocket is also consistent with the ability of A2U to bind a number of environmentally important chemicals found in unleaded petrol and shown to induce A2U nephropathy in rats³.

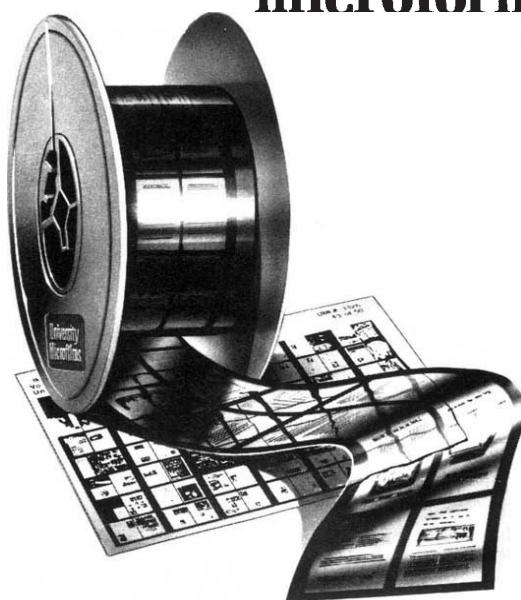
Our results give persuasive support to the proposition that A2U and MUP function as pheromone transport proteins in male rodent urine and will help in understanding the molecular basis of the modification of behaviour in rodents. They also provide detailed information on the structural basis for ligand binding by these proteins, which may be important in the metabolism of xenobiotic compounds in the rat and the metabolism of fatty acids in the kidney¹⁰. □

Received 5 May; accepted 25 September 1992.

1. Cowley, J. J. *Biological Determinants of Sexual Behaviour* (ed. Hutchinson, J. B.) 87–125 (Wiley, Chichester, 1978).
2. Cavaggioni, A., Sorbi, R. T., Keen, J. N., Pappin, D. J. C. & Findlay, J. B. C. *FEBS Lett.* **212**, 225–228 (1987).
3. Borghoff, S. J., Short, B. G. & Swenberg, J. A. *Rev. Pharmac. Toxicol.* **30**, 349–367 (1990).
4. Flower, D. R., North, A. C. T. & Attwood, T. K. *Biochim. biophys. Res. Commun.* **180**, 69–74 (1991).
5. Tirindelli, R., Keen, J. N., Cavaggioni, A., Eliopoulos, E. E. & Findlay, J. B. C. *Eur. J. Biochem.* **185**, 569–572 (1989).
6. Sivaprasadarao, A. & Findlay, J. B. C. *Biochem. J.* **255**, 561–569 (1988).
7. Böcskei, Z. et al. *J. molec. Biol.* **218**, 699–701 (1991).
8. Bacchini, A., Gaetani, E. & Cavaggioni, A. *Experientia* (in the press).
9. Cavaggioni, A., Findlay, J. B. C. & Tirindelli, R. *Comp. Biochem. Physiol.* **96b**, 513–520 (1990).
10. Kimura, H. et al. *J. Biol. Chem.* **266**, 5963–5972 (1991).
11. Kabsch, W. *J. appl. Crystallogr.* **21**, 67–71 (1988).
12. Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
13. Cherfils, J. & Dumas, C. *J. appl. Crystallogr.* **21**, 985–987 (1988).
14. Brunger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
15. Hendrickson, W. A. *Meth. Enzym.* **115**, 252–270 (1985).
16. Fitzgerald, P. M. D. *J. appl. Crystallogr.* **21**, 273–288 (1988).
17. Jones, T. A. *Meth. Enzym.* **115**, 157–171 (1985).
18. Cowan, S. W., Newcomer, M. E. & Jones, T. A. *Proteins* **8**, 44–61 (1990).
19. Papiz, M. A. et al. *Nature* **324**, 383–385 (1986).
20. Clark, A. J., Ghazal, P., Bingham, R. W., Barrett, P. & Bishop, J. O. *EMBO J.* **4**, 3159–3159 (1985).
21. Clark, A. J., Clissold, P. M., Shaw, R. A., Beattie, P. & Bishop, J. O. *EMBO J.* **3**, 1045–1052 (1984).
22. Ichihoshi, Y., Endo, H. & Yamamoto, M. *Biochim. biophys. Acta* **910**, 43–51 (1987).
23. Parry-Smith, D. J. & Attwood, T. K. *CABIOS* **7**, 233–235 (1991).
24. North, A. C. T. *Biochem. Soc. Symp.* **57**, 35–48 (1991).
25. Holden, H. M., Rypniewsky, W. R., Law, J. H. & Rayment, I. *EMBO J.* **6**, 1565–1570 (1987).
26. Huber, R. et al. *J. molec. Biol.* **195**, 423–434 (1987).
27. Flower, D. R. *J. molec. Graph* **9**, 257–258 (1991).

ACKNOWLEDGEMENTS. We thank N. Ito for discussion and the SERC for financial support. D.R.F. holds a Wellcome Prize studentship and S.E.V.P. holds an SERC Senior Fellowship. Coordinates will be deposited in the Brookhaven Protein Data Bank.

nature is available in microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone (____) _____

University
Microfilms
International

Detection of human DNA–carcinogen adducts

Fred F. Kadlubar

The role of specific chemicals in the aetiology of human cancer can now be assessed using ^{32}P -postlabelling in combination with corroborative biomonitoring techniques.

THIRTY years after Rachel Carson¹ raised the awareness of the general public to the potential dangers of environmental toxins, the application of multiple biomonitoring techniques has finally provided hard evidence for human exposure to genotoxic chemicals. At a recent International Meeting* focused primarily on ^{32}P -postlabelling methods, the detection of DNA–carcinogen adducts in a variety of human tissues was reported in association with occupational, dietary, tobacco and drug exposures. Most importantly, ^{32}P -postlabelling/thin-layer chromatography methods were corroborated by mass spectrometry, immunoaffinity and high-performance liquid chromatography, and fluorescence techniques and provided better estimates of DNA adduct levels during acute and chronic exposures. These efforts also led to the identification of specific DNA-bound chemicals such as 4-aminobiphenyl and benzo[*a*]pyrene, which are known carcinogenic components of cigarette smoke and other combustion sources. The upshot of these developments may be the beginning of a new approach to human risk assessment based on DNA–carcinogen adducts as molecular biomarkers that are predictive of clinical disease.

DNA–carcinogen adduct formation

DNA–carcinogen adducts are formed during the biotransformation of chemical carcinogens by drug-metabolizing enzymes to reactive intermediates that are electrophilic and bind covalently to DNA². The DNA-bound chemicals or adducts can be removed by DNA repair processes or by cell death, but upon chronic exposures often reach steady-state levels in carcinogen-target tissues. During cell replication, the DNA adducts can result directly in mutations in genes that control cell growth and lead to neoplasia³. Although DNA adducts of different carcinogens exhibit widely different mutation efficiencies, the steady-state levels of a specific DNA–carcinogen adduct in target tissues during chronic exposure appear to be dose-related and to be generally predictive of tumour incidence across species⁴. Thus, the accurate estimation and identification of human DNA–carcinogen adducts are expected to be predictive of human disease risk and pilot studies re-

ported at the Meeting provided initial support for this hypothesis.

Methods development

In experimental animal studies, the quantitation of DNA adducts had usually required the use of highly radioactive chemical carcinogens. However, a major breakthrough in detection methods occurred in the early 1980s with the development of the ^{32}P -postlabelling techniques^{5–8}. The method is based on enzymic hydrolysis of non-radioactive carcinogen-modified DNA to 3'-nucleotides, subsequent [5'- ^{32}P]-phosphorylation by ^{32}P -ATP (3,000–7,000 Ci mmol⁻¹) and polynucleotide kinase, and chromatographic separation of nucleotide–carcinogen adducts from normal nucleotides. The result was an assay of remarkable sensitivity, with detection limits of about one adduct per 10⁹ normal nucleotides (≈ 3 adducts per genome).

Quantifying adduct levels

To date, over 60 laboratories have used and customized these methods and most of these practitioners met together for the first time in June to discuss their art. A unique aspect of the Meeting was an international effort by 15 different laboratories to quantify DNA–carcinogen adduct levels in chemically-modified DNA standards and in human tissue samples. Although the results were generally comparable between most laboratories, an unexpected outcome was the apparent underestimation of total adduct levels by this method. The implication, of course, is that DNA adduct levels in human tissues may actually be higher than they are now estimated to be using ^{32}P -postlabelling. Moreover, it became clear that, for quantitative purposes, adduct stability during isolation and storage, adduct recovery in the chromatography systems used, and the efficiency of enzymatic [5'- ^{32}P]-phosphorylation must be determined. This will require the further availability of synthetic DNA–carcinogen adduct standards that are representative of different chemical classes and the development of additional corroborative/alternative methods. With these caveats, it is expected that accurate measurements can be achieved, that additional DNA–carcinogen adducts can be identified in human tissues and that environ-

mental monitoring for future health effects can be accomplished. But even now, the combined application of such standardized methods has already provided strong evidence for the presence of DNA adducts derived from aflatoxin, aromatic and heterocyclic amines, polycyclic aromatic hydrocarbons, nitrosamines and chemotherapeutic drugs in humans.

Further developments

Another significant development was the wide application of these methods to the analysis of both structurally simple (low molecular weight) and bulky carcinogens in complex mixtures, as well as DNA adducts derived from irradiation. Aquatic organisms and plants were also sampled and found to contain DNA adducts that could be attributed to environmental pollution. Finally, one of the more exciting findings was that paraffin-embedded tissues, if fixed less than 48 hours in formalin, are quite suitable for ^{32}P -postlabelling analyses. This should allow for many new studies on DNA available for isolation from human tissue archives stored worldwide and should provide a wealth of new information on the role of chemical carcinogens in human cancer. □

**International Meeting on Postlabelling Methods for the Detection of DNA Adducts, International Agency for Research on Cancer, Lyon, France, 24–27 June 1992.*

Fred F. Kadlubar is associate director for research at the National Center for Toxicological Research, US Food and Drug Administration, Jefferson, Arkansas 72079, USA. For more information, fill in reader service number 100.

1. Carson, R. *Silent Spring* (Fawcett, Greenwich, Connecticut, 1962).
2. Miller, J. A. and Miller, E. C. *Environ. Hlth Perspect.* **49**, 3–12 (1983).
3. Vogelstein, B. et al. *New Engl. J. Med.* **319**, 525–532 (1988).
4. Gaylor, D. W., Beland, F. A. and Kadlubar, F. F. *Environ. Hlth Perspect.* (in the press).
5. Randerath, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6126–6129 (1981).
6. Gupta, R. C. et al. *Carcinogenesis* **3**, 1081–1092 (1982).
7. Gupta, R. C. *Cancer Res.* **45**, 5656–5682 (1985).
8. Reddy, M. V. and Randerath, K. *Carcinogenesis* **7**, 1543–1551 (1986).

Cell biology: back to basics

A chamber for assessing the invasive properties of cells *in vitro*, tritium-labelled AZT-triphosphate and a gel microdrop maker are a selection of this week's new products in the scientific marketplace.

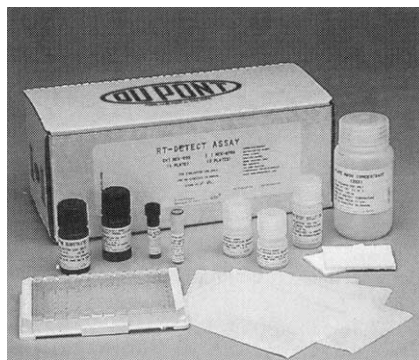
THE Expert-Center for Taxonomic Identification at the University of Amsterdam has introduced the first in a series of scientific **CD-ROM disks** that are suitable for the identification of organisms, the study of biodiversity, the preservation of the environment and education (*Reader Service No. 101*). Produced at the Institute of Marine Research in Bergen, Norway, 'Linnaeus Protist' is a collection of information on the identification and ecology of North Atlantic algae. Information provided by more than a dozen prominent North European scientists is included on 312 algal species, many of which also occur in other parts of the world. The information includes test descriptions, literature references, onscreen micrographs of photographic quality, methods suitable for algal groups, distribution maps of toxic and bloom-forming organisms, an introductory section on protist groups and life cycles and a glossary of scientific terms used in the program. Linnaeus Protist requires a Macintosh Plus or higher computer with at least 1 MB of RAM and HyperCard 2.0 or higher. A colour monitor is suggested for the viewing of micrographs. Linnaeus is compatible with System 7. Through special funding from UNESCO, the Center is able to make Linnaeus Protist and the accompanying manual available to researchers for the shipping and handling charge of \$32 (US). Work is in progress on CDs covering other groups of organisms and environmental areas.

Collaborative Biomedical Products/Becton Dickinson now offers a ready-to-use research tool that is designed to facilitate the **study of cell invasion *in vitro*** (*Reader Service No. 102*). The Biocoat Matrigel Brand Invasion Chamber consists of a Falcon cell culture insert containing an 8- μ m pore PET membrane. The product is supplied as two 24-well culture plates, each of which contains 12 invasion chambers coated precisely with 100 μ g per cm^2 of

Matrigel basement membrane matrix. The company says that these chambers are suitable for a variety of invasion assay protocols: the invasion of cells through basement membrane matrix *in vitro* has been shown to correlate with Matrigel invasion *in vivo*. It has been reported that such invasion chambers can be used to determine the effects of a drug such as Taxol on metastatic activity, as well as to identify cellular mechanisms of normally invasive cells such as cytotrophoblasts. The invasion chamber membrane can be removed with a scalpel and can be processed for light and electron microscopy image analysis, cell counts or for counting radioactivity.

Reagents

Du Pont hopes to make retrovirus research easier and safer with the introduction of its **RT-Detect reverse transcriptase assay**—



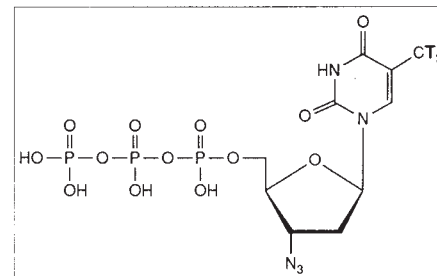
New RT-Detect assay from Du Pont.

a nonradioactive assay that is designed to eliminate many of the problems associated with conventional reverse transcriptase methodology (*Reader Service No. 103*). A measurement of reverse transcriptase activity is often used as a method of detecting human immunodeficiency virus and other retroviruses in tissue culture supernatants. The standard method measures the incorporation of tritium-labelled thymidine into acid-precipitable material in the presence of poly(rA) and an oligo(dT) primer, followed by centrifugation or filtration and scintillation counting. Because contaminating cellular polymerases can generate a positive result, Du Pont says that it is necessary to run each sample in the ^3H -based method with poly(A) and poly(dA) templates to show the presence of a reverse transcriptase in the sample. The complex and tedious manipulations required to separate the incorporated from unincorporated label limit the number of samples that can be processed at one time to about 35. RT-Detect is designed to provide the same sensitivity but

without the need for radioactivity. And, in addition, because it uses a specific heteropolymeric RNA template, contaminating cellular polymerases are inactive and a separate reaction control for their presence is not necessary. Du Pont says that the entire assay, from addition of the sample to calculation of the results, can be completed in just over four hours.

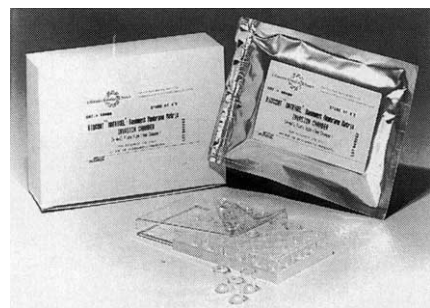
Six sets of **quality-control standards for cystic fibrosis testing** are now available from BioVentures (*Reader Service No. 104*). CustomCuts are designed for use as either external or, preferably, as internal controls for determining polymerase chain reaction (PCR) restriction fragment length polymorphisms associated with certain genetic disorders. When digested with the designated restriction enzyme, CustomCuts provide information with respect to enzymatic activity and the adequacy of digestion. CustomCuts also provide an estimate of DNA size. Undigested and digested CustomCuts are designed to provide bands distinct from those present in the PCR products. BioVentures says that they are particularly useful for interpreting the banding pattern given by heterozygotes in which half the products contain the site and the other half lack the site.

Since March 1987 when AZT (azidothymidine or zidovudine) was approved for clinical use in the treatment of AIDS by the US Food and Drug Administration, research into the virus-inhibiting properties of AZT and related azido compounds has continued apace. With this in view, Moravsek Biochemicals is offering researchers **tritium-labelled AZT-triphosphate**



Tritium-labelled AZT-triphosphate.

(*Reader Service No. 105*). Tritium-labelled AZT-triphosphate has a specific activity of 3–20 Ci mmol^{-1} and is available as 0.25 mCi and 1 mCi. AZT-triphosphate, [methyl- ^3H]- is shipped in an ethanol:water (1:1) solution on dry ice. Tritium-labelled AZT, AZT monophosphate as well as nonlabelled versions of AZT-
NATURE · VOL 360 · 12 NOVEMBER 1992

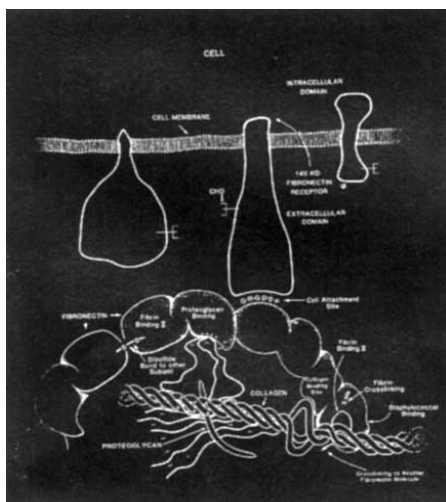


Tool for studying cell invasion *in vitro*.

monophosphate and AZT-triphosphate are also available.

Bafilomycin A₁, which is available from Kamiya Biomedical, is a **macrolide antibiotic** isolated from cultures of *Streptomyces griseus* (Reader Service No. 106). Kamiya says that Bafilomycin A₁ is highly specific, inhibiting vacuolar-type H⁺-ATPase without affecting other ATPases. Bafilomycin A₁ exhibits activity against Gram-positive bacteria and fungi and appears to inhibit several cellular processes, including lysosomal acidification and L-glutamate uptake, presumably through its action on H⁺-ATPase. The company expects it to be a useful research tool in the study of several areas of cell regulation, including protein degradation and proton transport and for distinguishing between different types of ATPases.

Telios Pharmaceuticals has a wide selection of **extracellular matrix proteins** for research use — details of which are contained within the company's 1992–1993 research products catalogue (Reader Ser-



Purified proteins for extracellular matrix research — see Telios.

vice No. 107). The line-up includes fibronectins, human fibronectin proteolytic fragments, human collagens, vitronectins, laminins and merosins. Telios says that these proteins are highly purified and suitable for use in assays such as ELISAs, immunoblotting and cell attachment studies. The company is also offering a line of related antibodies, cDNA probes and peptides.

Latest in labware

This month, One Cell Systems, a company focused on developing clinical diagnostic and research products based on its proprietary **gel microdrop technology**, began selling its CellSys 100 gel microdrop (GMD) maker (Reader Service No. 108). The \$7,500 (US) CellSys 100 system includes the GMD maker, an ice-water bath container, a sidearm

flask and an autoclavable (120 °C) stainless steel shaft and blade assembly. The system is an emulsifier designed to maximize cell encapsulation while preserving cell integrity. High rotation speeds and the special blade configuration allow a choice of 10-μ to 200-μ diameter drop size. The maker is equipped with a moveable stage, allowing immersion of the oil–agarose emulsion in an ice-water bath while maintaining the proper shear force for uniform preparation of solidified GMDs. The company's GMD technology permits the analysis and separation of cells based on cell function such as growth, drug susceptibility, protein secretion, enzyme activity or metabolite production, or based on cell composition such as surface markers, internal proteins or nucleic acid sequences. Three basic steps are common to all applications: 1) forming GMDs from a cell suspension, creating microenvironments that can be manipulated, 2) testing single cells and microcolonies within many GMDs simultaneously and 3) analysing and isolating GMDs of interest using



One Cell Systems' gel microdrop maker.

physical methods such as flow cytometry. Gel microdrops are prepared by dispersing a mixture of molten agarose and cells into an excess of oil to form an emulsion. As the emulsion cools, the drops gel. In this form, One Cell Systems says that the GMDs are physically distinct and can be handled like cells: they can be centrifuged, collected and dispensed by pipette and incubated. The company also offers encapsulation, secretion and hybridoma selection assays to accompany the CellSys 100 system.

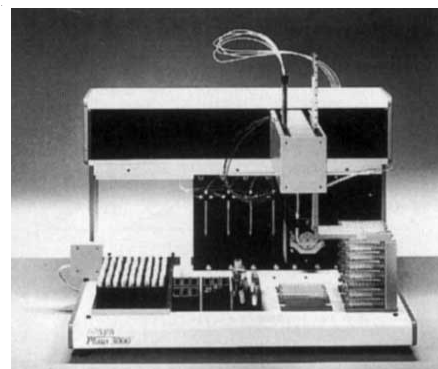
With the recent acquisition by Endo-tronics of the Opticell product line from Charles River Laboratories, the company is now able to offer a range of fully **automated cell culture systems** (Reader Ser-



Opticell cell culture system.

vice No. 109). Opticell systems (5200R, 5200E and 5300E) are designed for the immobilization, growth and long-term maintenance (continuous perfusion) of both suspension and anchorage-dependent cells, including vaccine applications. Opticore bioreactors of varying size, which are based on a patented ceramic matrix, provide users with flexibility and scalability. Operating parameters such as medium flow, pH, dissolved oxygen and temperature are continuously monitored and controlled, with the oxygen consumption rate calculated online as medium perfuses directly across the culture.

Rosys, a manufacturer of robotic liquid handling instruments, has developed a new **robotic system for microplate preparation and analysis** (Reader Service No. 110). The Plato 3000 system is capable of automating liquid handling, plate washing, read-



Automate your lab with the Plato 3000.

ing and incubation. Plates are housed in plate stackers to minimize the size of the instrument without affecting capacity. A robotic arm transports the plates between the stackers, the work surface and other modules incorporated in the system. The liquid handling module is multitipped with a choice of washable or disposable tips. Plato 3000 also includes modules for positive sample identification, a stacker/incubator module for up to 16 plates and plate washing and plate reading modules. PC-driven software is available

ADVERTISEMENTS

Teflon^{*} Coated and Bare Wire from MEDWIRE



Medwire provides the highest quality platinum, platinum alloy, silver and stainless steel wire in both single and multistrand configurations.

For prompt delivery or additional information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg. TM of E. I. DuPont & Co

Reader Service No.25

WHY PAY MORE ?

MINI-ELECTROPHORESIS

Model 100 for 92 x 100mm gel in 100ml buffer
or

Model 101 for 80 x 100mm gel in 250ml buffer

Including Combs for 6, 8, 12 & 16 wells, UV-transparent gel support, platinum electrodes and lid-mounted cables for safety & Swiss quality from DANAPHOR^R for US \$ 140.-(Air-freight & Insurance extra).

TECTATE S. A.

2 place Grand St.-Jean,
1003 Lausanne, Switzerland.
for invoice Fax: 0041-21-311 5783.

Reader Service No.40

HPLC, TLC, GC

RADIOACTIVITY
+
CHROMATOGRAPHY



raytest

raytest • Isotopenmeßgeräte GmbH
D-7541 Straubenhardt 1 • Benzstraße 4
Tel. 07082 / 6 04 07 • Fax 07082 / 2 08 13

...WE LISTEN TO OUR CUSTOMERS

Reader Service No.29



Seymour Biotech Ltd.

CUSTOM SERVICES

- **OLIGONUCLEOTIDE SYNTHESIS**
High quality PROBES, PRIMERS, ANTISENSE OLIGOS!
Fast-Efficient Service offering loads of modifications
BIOTIN, FLUORESCENCE, 5' & 3' labeling,
Phosphorothioates Affinity Support/Oligos, etc.
- **MULTIPLE PEPTIDE SYNTHESIS**
Epitope Mapping & Screening! We now supply 5mg
quantities of high purity custom peptides up to 15
residues in length, using Fmoc chemistry. With D or L
amino acids, these are supplied lyophilized in 4ml
vials, cleaved from solid phase & ready for use

Tel: 0562 825286 Fax: 0562 825284

Unit 23, Huddersfield Enterprise Centre
Worcester Road, Kildersleepe, DY10 1HY

Reader Service No.26

with application programs for routine assays or more complex routines.

Manufactured by Professional Diagnostic and distributed by B/T SciTech, the new **anaerobic tissue processor** is designed to eliminate many of the problems associated with conventional homogenizers, namely the introduction of air and heat to the homogenized sample (*Reader Service No. 111*). The device can be used in a cold-room



Cold-room compatible anaerobic tissue processor — see B/T SciTech.

and offers the user a touchpad control panel on the motorized base and a choice of six homogenization chambers ranging in size from 5–250 ml. Speeds can be varied from between 1,000 and 28,000 r.p.m. and processing times can be set anywhere from 1–99 s. B/T SciTech says that both soft (for example, spleen) and solid (for example, muscle) tissue can be processed faster and with increased cell yield than with more conventional methods.

Histology

HistoChoice is a new, **nonhazardous tissue fixative** from AMRESKO, which does not contain formaldehyde, mercury, picric acid or any other toxic materials (*Reader Service No. 112*). This non-alcohol based fixative is odourless and can be safely poured down drains. HistoChoice is suitable for all tissue types and compatible with



Fix it with HistoChoice fixative.

all manual and automated tissue processing equipment and procedures. In addition, AMRESKO says that as it is osmotically balanced, cellular and tissue distortions should be eliminated. Standard haematoxylin and eosin Y stains show sharp and vibrant staining of nuclear and cytoplasmic components and enhanced preservation of morphologic detail. AMRESKO says that overfixation is not a problem and cross-linking of antigen binding sites is not observed.

With a throughput of 160 slides per hour, Shandon Scientific says that its Consul **automatic slide coverslipper** enhances consistency and frees laboratory staff from the time-consuming and laborious task of manually coverslipping slide-mounted specimens (*Reader Service No. 113*). Up to 40 stained specimens (sections or smears) can be transferred and loaded directly into



Shandon's automatic coverslipper.

the Consul for coverslipping using the special slide clips and rack provided. Standard glass slides (76 x 26 mm) are covered with routinely-accepted coverslips (40, 50 or 60 mm) using commercially available mountants. The 'roll on' pressure pad is designed to ensure that each coverslip achieves smooth, even adhesion each time. The volume of mountant can be adjusted using the nine available settings. Shandon says that a rack of 40 slides can be fitted with coverslips in 15 min. Coverslips that are angled incorrectly or stuck together are detected by optical sensors. A built-in exhaust fan and filter minimizes operator exposure to xylene fumes. Additional protection is provided by an alarm that is triggered by opening the protective hood when the instrument is in operation.

Antibodies

Endoglin is an endothelial homodimeric membrane antigen containing the RGD (Arg-Gly-Asp) recognition motif for adhesion receptors of the integrin family. The glycoprotein (170–180 kDa) is present on endothelial cells of capillaries, arterioles and venules in a variety of tissues and at low levels on acute lymphoblastic and myelocytic leukaemia cells. The **monoclonal antibody to endoglin**, which is available

NATURE • VOL 360 • 12 NOVEMBER 1992

ADVERTISEMENTS

from Cymbus Bioscience, has been used to stain human spleen, tonsil, lymph node and liver cells (*Reader Service No. 114*). The company also offers monoclonal antibodies that will stain all aspects of the leukocyte common antigen in formalin-fixed, paraffin wax-embedded tissues and a monoclonal antibody to major basic protein. The latter should be a useful pan-eosinophilic marker as it has been shown to stain all human eosinophils — whatever their degree of activation.

PharMingen announces the availability of a monoclonal antibody pair optimized for the **quantitation of mouse interleukin-2 (IL-2)** (*Reader Service No. 115*). This monoclonal antibody pair allows researchers to assay IL-2 using a fast, easy and economical ELISA format that provides high sensitivity and reproducibility, PharMingen says. The assay is intended to offer significant advances over the time-consuming and labour-intensive IL-2 bioassay, which requires maintenance of cell lines and the use of radioisotopes. Each monoclonal antibody pair provides enough reagent to assay about forty 96-well plates.

BABCO is supplying the monoclonal antibody 12CA5 (anti-influenza haemagglutinin HA1), which is widely used in **epitope tagging and protein surveillance studies** (*Reader Service No. 116*). Fusion proteins are expressed from a vector encoding a sequence specific for a nine-amino acid peptide epitope at the terminus of the protein of interest. BABCO says that the specificity and affinity of the 12CA5 antibody to this nonapeptide epitope makes it suitable for use in immunofluorescence immunoprecipitation, western blotting, quantitation and affinity purification of recombinant fusion proteins. The 12CA5 antibody, which was developed at the Scripps Research Institute, is supplied in a purified form or in ascites fluid. The nine-amino acid peptide is also available.

Two new **monoclonal cell lines to human mitochondrial heat-shock protein 60 (HSP 60)** are now available from Advanced Biotechnologies (*Reader Service No. 117*). The monoclonals are supplied purified by either Protein A or gel filtration and have IgM and IgG2a antibody isotypes, respectively. Both antibodies have been used to coat sample wells in ELISA-based capture assays, in western blots after biotin labelling and for immunocytochemical studies on fixed cells. HSP 60 represents the human homologue of the mycobacterial HSP 65 protein and is implicated as a target autoantigen in a variety of autoimmune diseases, including rheumatoid arthritis and multiple sclerosis, as well as schizophrenia. HSP 60 also acts as a chaperonin in the process of protein folding, is constitutively expressed in normal

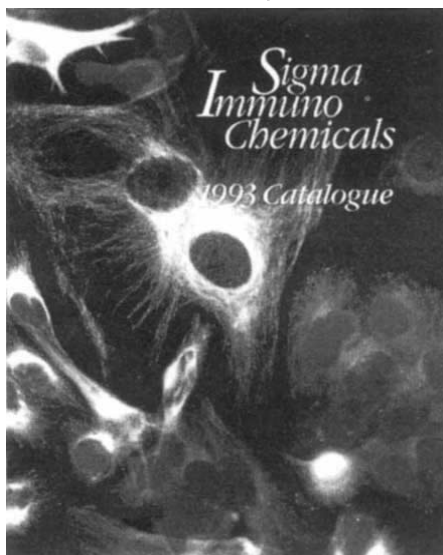
human synovium and has recently been localized to cell mitochondria.

Cambridge Research Biochemicals offers two new antibodies to complement its range of **oncoprotein markers** — a sheep polyclonal antibody to *fos* and a rabbit polyclonal antibody to *c-erbB2* (*Reader Service No. 118*). Both antibodies have been extensively characterized for use in western blotting, immunocytochemistry and immunoprecipitation. The immunizing peptides are also available for each of these antibodies.

According to BioGenex, its monoclonal antibody, QBEnd/10, which reacts with the CD34 antigen in human endothelial and haematopoietic cells, should provide a more specific method of **evaluating vascularization** — an important focus of tumour evaluation (*Reader Service No. 119*). It stains positively in routinely fixed paraffin sections from a variety of vascular and lymphatic tumours (*Histopath. 17*, 237–242; 1990). In a study of malignant vascular tumours of the liver by Anthony (*J. clin. Path. 44*, 29–32; 1991), it was found that QBEnd/10 was more sensitive and specific than antibodies to *Ulex europaeus* 1 agglutinin and Factor VIII-related antigen. This monoclonal antibody reacts specifically with endothelium in formalin-fixed, paraffin-embedded or frozen tissue sections. It is available in concentrated form or as a ready-to-use reagent that is optimized for use with the Super Sensitive and High Performance biotin-streptavidin detection systems.

Off the shelf

Copies of the new 1993 **immunochemicals catalogue** from Sigma are now available (*Reader Service No. 120*). The catalogue lists over 1,400 immunochemicals, together with technical data and information on applications. Among the new product introductions are Quantum Red direct conjugates, new CD markers, Sigma Fast substrate



Sigma's 1993 immunochemicals catalogue.

Zeinh 1D/2D/DNA/TLC

HAND-HELD SCANNING DENSITOMETER

To quantify:

Gels, blots, TLC, X-Ray film, Photo. 1-D pkg price \$995

8 bits, 60 um resolution, 0-3 OD, sample 11x40 cm, IBM or clones

Options: UV, 2-D, DNA, ELISA, colony counting..... \$995 each

Universal software: for any scanner. Main 1D, 2D, DNA, \$995 each

Representatives needed, 30% discount of USA price list.

Do-it-all, Departmental, Laser + Video

Micro Densitometry Station

\$ 16,990

Hi-Speed Soft Laser

BIOMED INSTRUMENTS INC. 1020 S Raymond Av. #B, Fullerton CA 92631. Pnn: (800) 927-2290, (714) 870-0290, Fx: 17148706385

Reader Service No.22

Professional Custom Services

DNA SYNTHESIS
PROTEIN SEQUENCING
PEPTIDE SYNTHESIS

Rapid Service • Guaranteed Quality • Ample Quantity
Competitive Prices • Free Quotations
World-wide Delivery

V.G. Services 1200 Wonderland Rd S. Bldg #7 Unit #2
London, Ontario Canada N6L 1A8

CALL FAX: (519) 652-3758 ALT. FAX: (519) 652 0053

Reader Service No.32

\$3.50/Base

Custom DNA Synthesis

No Set-up Fee,
Shipped in 24 hours.

1-800-788-4362
1-800-786-4362 Fax

KEYSTONE Laboratories, Inc.

Reader Service No.33

Oligos Etc. Inc.

DNA \$4.50/base
S-DNA \$8/base
Methyl Phosphonate DNA \$8/base
RNA \$8/base
S-RNA \$12/base
O-Methyl RNA \$15/base
Chimerics, Biotin
Modifiers, Fluorescent Tags etc.
No Set-Up Charges

800-888-2358/fax 800-869-0813
Intl.fax 203-438-3637

Reader Service No.30

ADVERTISEMENTS

Reader Service No.56

Highest Quality Oligonucleotides

Synthesised, purified and shipped in 3 days

£2.75 per base

(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit

King's College School of Medicine & Dentistry

Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No.15

Synthetic Oligonucleotides

Rapid Quality Worldwide Service

£2.50 per base, no setup!

option of H.P.L.C. purification
biotinylation, fluorescent labelling
RNA Synthesis

phosphothioates, many other modifications

OSWEL DNA SERVICE

Department of Chemistry, University of Edinburgh
Kings Buildings, West Mains Road,
EDINBURGH EH9 3JJ

Tel: 031 650 4792 Fax: 031 662 4054

Reader Service No.34

MICROMANIPULATORS

XENOPUS OOCYTE INJECTION

- Mk 1 "Pantograph" Micromanipulator
- Manual & Automatic Injectors
- Microscopes • Pipette Pullers

SINGER

INSTRUMENTS

"A responsibility to Science"

Call or Fax now for further information

Tel: +44 984 40226 Fax: +44 984 41166

Singer Instrument Co Ltd, Roadwater, Watchet, Somerset, England, TA23 0QL

PACEMAKERSM

The multiple peptide synthesis and conjugation service

- High purity (up to 95%).
- Quantities for a range of uses.
- Single or multiple peptides.
- HPLC trace with each peptide.
- Terminal modification at no extra charge.
- Conjugation to carrier proteins.

from **£80 per peptide**

AFFINITY Research Products Ltd.

GPT Business Park, Technology Drive,
Nottingham, NG9 2ND, UK.

Tel: +44/0 602 436100; Fax: +44/0 602 436300

tablets, new cytokines and growth factors, rapid immunohistological staining kits and new conjugates. Over 200 new products can be found in the company's 1993 Cell Culture Catalogue. The catalogue lists a range of products for the detection and elimination of mycoplasma infections. The company's mycoplasma agar and broth can be used to propagate mycoplasma in cultures infected at a low level or with a fastidious strain. In addition, bisBenzimide and 4'-6-diamidino-2-phenylindole are available for the direct detection of mycoplasma by DNA fluorochrome staining: mycoplasma-infected cultures will exhibit small, uniformly shaped fluorescent bodies in the extranuclear and intranuclear spaces. Infected cultures can be treated with the fluoroquinolone antibiotic, ciprofloxacin, which acts by inhibiting the mycoplasma DNA gyrase. Sigma recommends the treatment of infected cultures with ciprofloxacin at a concentration of 10 mg l⁻¹ of culture medium for 12 days. Ciprofloxacin is effective against at least seven common mycoplasma species and, to date, has not been shown to induce resistance.

American Type Culture Collection has released its **culture catalogues on computer diskettes** (IBM PC or compatible) (Reader Service No. 121). These include bacteria and bacteriophages; cell lines and hybridomas; filamentous fungi; protists (algae and protozoa); ATCC's recombinant DNA materials; ATCC's/National Institutes of Health's repository of human and mouse DNA probes and libraries; animal viruses and antisera; plant viruses and antisera; and yeasts. The catalogue diskette set is comprised of eight sections which may be purchased individually or in any combination. The price per section is \$20 or \$25 (US), depending on the section. ATCC offers a 20 per cent discount when orders for two or more sections are placed at one time.

Word perfect

EVEN the simplest technical words such as borane or nicotinamide are apt to confound most spell-checker programs. Researchers, therefore, may be interested in two new programs from European Scientific Software — BioWords and ChemWords — that are designed to increase the scientific vocabulary of your word processor's spell checker (Reader Service No. 122). BioWords and ChemWords dictionaries are available for both Macintosh or IBM PCs or compatible computers. Each dictionary contains about 30,000 words relating to either biology or chemistry. So whether you write technical reports as a profession or only occasionally, ESC says that the programs should prove to be an invaluable time-saver. The PC version is compatible only with Word Perfect, whereas the version for the Macintosh is compatible with most Macintosh word-processing systems.

Gibco BRL has good news for cell biologists — the company is offering a free applications booklet on how to **separate cell organelles** more efficiently (Reader Service No. 123). The booklet provides



All you need to know about separating cell organelles.

protocols for and discusses the use of Nycomed density gradient media. Sections outline the preparation and fractionation of nuclei, mitochondria, lysosomes and peroxisomes. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIABORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.23